REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (*Leave blank*)

2. REPORT DATE
Jan 95

3. REPORT TYPE AND DATES COVERED

4. TITLE AND SUBTITLE
Analysis of Group Performance in Visual Signal Detection

5. FUNDING NUMBERS

6. AUTHOR(S)
Christopher J. Hays

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
AFIT Students Attending:
University of Florida

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT/CI/CIA
94-145

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
DEPRTMENT OF THE AIR FORCE
AFIT/CI
2950 P STREET
WRIGHT-PATTERSON AFB OH 45433-7765

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for Public Release IAW 190-1
Distribution Unlimited
MICHAEL M. BRICKER, SMSgt, USAF
Chief Administration

12b. DISTRIBUTION CODE

13. ABSTRACT (*Maximum 200 words*)

19950103 036

14. SUBJECT TERMS

15. NUMBER OF PAGES
103

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT

18. SECURITY CLASSIFICATION OF THIS PAGE

19. SECURITY CLASSIFICATION OF ABSTRACT

20. LIMITATION OF ABSTRACT

ANALYSIS OF GROUP PERFORMANCE IN VISUAL SIGNAL DETECTION

By

CHRISTOPHER J. HAYS

A THESIS PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

UNIVERSITY OF FLORIDA

1995

ACKNOWLEDGEMENTS

To begin with, I would like to thank my caring and helpful wife, Tiffany. When I got discouraged or needed support, she was always there. Toki Sadralodabai is a great friend and has helped me succeed in finishing this program and has shown me by example how to succeed. I owe her a great deal of thanks.

I would also like to thank my advisor, Dr. Sorkin, for all his help in preparing my thesis, helpful advice, and the opportunity to further my education at the University of Florida. I would like to acknowledge and thank my committee members, Dr. David Green and Dr. James Shepperd, for their time and advice regarding this study.

An additional acknowledgement goes to Josh Fryman for his help in programming and Dr. Kourosh Saberi with his help on the computer simulations.

This research was partially supported by grants from the Air Force Office of Scientific Research.

TABLE OF CONTENTS

LIST OF TABLES

Table | Page

LIST OF FIGURES

Figure		Page

Figure | Page

Abstract of Thesis Presented to the Graduate
School of the University of Florida in Partial Fulfillment
of the Requirements for the Degree of Master of Science

ANALYSIS OF GROUP PERFORMANCE IN
VISUAL SIGNAL DETECTION

By

Christopher J. Hays

May, 1995

Chairman: Robert D. Sorkin
Major Department: Psychology

This study analyzed how a group's decision making performance compared to the predicted performance of a statistically "optimal" group on a visual signal detection task. The task was to decide, either individually or as a group, whether a stimulus display screen (composed of nine independent information sources) was representative of one of two possible normal distributions (signal or noise). Eight subjects were tested individually and as members of different sized groups. Sorkin and Dai's 1994 model of the Ideal Group was used to analyze the accuracy of group performance (d'_{grp}). Consistent with the model of the Ideal Group, d'_{grp} increased with the number of group members, but decreased with inter-member correlation and task difficulty.

The group decision process was assumed to be based on the group assigning different weights (emphasis) to each group members' input. Overall efficiency, a relative measure

of actual performance compared to ideal performance, was relatively high for two-member groups but decreased as the group grew larger. Group weighting strategies appeared to be somewhat variable for the condition in which all sources of information presented to the members were equally valid. However, when the sources of information were unequal, the group appropriately assigned more weight to the group members with the more reliable information. Some interesting social interaction phenomena were theorized to take place during the group decision process. Social loafing, coordination loss and outgoingness were examined as possible reasons for changes in group performance. Interestingly, the subjects received more weight when they acted as the spokesperson for the group compared to when they did not respond for the group.

Title of Master's Thesis:
Analysis Of Group Performance in Visual Signal Detection

Author:
Christopher J. Hays

1LT, United States Air Force

1995

103 pages

Master of Science

University of Florida

References (* Denotes importance):

Berg, B. G. (1990). Observer efficiency and weights in a multiple observation task. Journal of the Acoustical Society of America, 88, 149-158.

Durlach, N. I., Braida, L. D., & Ito, Y. (1986). Towards a model for discrimination of broadband signals. Journal of the Acoustical Society of America, 80, 60-72.

Green, D. M. (1994). Signal detection theory revisited. Unpublished manuscript.

* Green, D. M., & Swets, J. A. (1966). Signal detection theory. New York: Wiley.

* Grofman, B., Owen, G. & Feld, S. L. (1983). Thirteen theorems in search of the truth. Theory and Decision, 15, 261-278.

Harkins, S., & Jackson, J. (1985). The role of evaluation in eliminating social loafing. Personality and Social Psychology Bulletin, 11, 457-465.

Hinsz, V. B. (1990). Cognitive and consensus processes in group recognition memory performance. Journal of Personality and Social Psychology, 59, 705-718.

Latané, B., Williams, K. & Harkins, S. (1979). Many hands make light the work: The causes and consequences of social loafing. Journal of Personality and Social Psychology, 37, 822-832.

* Lutfi, R. A. (in press). A correlational method for deriving weights in multiple-observation tasks. Journal of the Acoustical Society of America.

* Shepperd, J. A. (1993). Productivity loss in performance groups: A motivation analysis. Psychological Bulletin, 113, 67-81.

Sniezek, J. A., & Henry, R. A. (1989). Accuracy and confidence in group judgment. Organizational Behavior and Human Decision Processes, 43, 2-28.

* Sorkin, R. D. & Dai, H. (1994). Signal detection analysis of the ideal group. Journal of Organizational Behavior and Human Decision Processes, 60, 1-13.

* Sorkin, R. D. & Montgomery, D. A. (1991). Effect of time compression and expansion on the discrimination of tonal patterns. Journal of Acoustical Society of America, 90, 846-857.

Vollrath, D. A., Sheppard, B. H., Hinsz, V. B., Davis, J. H. (1989). Memory performance by decision-making groups and individuals. Organizational Behavior and Human Decision Processes, 43, 289-300.

* Denotes reference of primary importance.

INTRODUCTION

Groups of all kinds frequently make important decisions based on visual sources of information. Often these decisions need to be made quickly and correctly. For instance, an airplane crew needs to decide if they are experiencing an actual engine problem or if the engine gauges are malfunctioning. Because group decisions can have important consequences, we are interested in determining how well groups perform compared to a statistically optimum group.

Researchers from diverse fields have studied separate variables and theories that affect group performance during the decision process. Grofman, Owen, and Feld (1983) studied the effect that number of group members had on group performance. Sniezeck (1990) analyzed how correlation between member's estimates of weather forecasts affected the group performance. Johnson and Torcivia (1967) investigated the impact of individual performance on group performance. The relative weights (emphasis) assigned to individual group members during the group interaction have also been studied (Sniezeck & Henry, 1988).

The aim of the current study is to determine how well an actual group performs a visual detection task compared to a model of the statistically optimal group, the Ideal Group (Sorkin & Dai, 1994). The equations describing ideal group

performance integrate several of the previously studied variables. This study will focus on determining how efficiently the actual groups weight the input of the members and how their performance differs from ideal. It is hoped that the comparison between the actual and ideal performance will lead to better predictions about group performance. This knowledge may then lead to more appropriate training in decision making groups.

Most of the prior research has not been concerned with the statistical analysis of group performance and their results have been described in general terms. For instance, the conclusion has been made that the group with three observers performed better than the group with two (Libby & Blushfield, 1978 from Grofman, Feld & Owen, 1984). But the exact increases and reasons for change in performance were not stated. Reliable predictions concerning group performance are difficult to make without mathematical analyses describing ideal performance.

Sorkin and Dai (1994) applied the Theory of Signal Detection (TSD, Green & Swets, 1966; Green, 1994) to derive the mathematical equations that describe the performance of an ideal group. Sorkin and Dai describe the performance of an ideal group in terms of the number of observers in the group (m), the correlation between the individual observers (r), the array of the individual observers performance prior to group interaction (d'_i), and the relative weights given to each observer's estimate (a_i). The current study will

examine how each of these variables actually affects group performance. Sorkin and Dai (1994) assume a linear group decision process based on free group interaction and the optimal weighting of group members' individual estimates in order to make a group decision. Although this is not the only possible group decision process (i.e., Delphi group, closed ballots, majority wins, plurality wins), it is the method that statistically provides the highest achievable group performance.

Group decision making processes have been studied for more than two centuries and are still not completely understood. The Condorcet Jury Theorem (Marquis de Condorcet, 1785) states that as long as each group member is correct at least 50 percent of the time (chance), group performance increases monotonically with the addition of group members. Therefore, by adding new group members, the group performance can rapidly approach an asymptote of perfect performance. However, the addition of members that perform worse than chance would result in the group performance actually decreasing (Grofman, Owen & Feld, 1983).

Grofman, Feld, and Owen (1984) also showed that the increase in group performance was affected by the difficulty of the task. If performance was less than 50 percent accurate, then the group performance would actually decrease with the addition of group members. Another theory that predicts group performance will be better than individual performance, although for a different reason, is called truth-wins.

Truth-Wins. The Best Member or Truth-Wins Model (Lorge & Solomon, 1955) contends that a group will perform better than an individual, because it is more likely that one of the members will recognize the correct decision and persuade the group to respond correctly. Although there are different names for this type of theory, e.g. Model A (Lorge & Solomon, 1955), Rational Model (Thomas & Fink, 1961), studies testing this theory (Lorge & Solomon, 1955; Thomas & Fink, 1961), have been inconclusive (Johnson & Torcivia, 1967). Hartwick et al. (1982) evaluated several studies and determined that this theory over-predicts group performance and is probably not how actual groups make decisions.

Error checking. Another possibility that has been suggested as to why groups typically outperform individuals is that groups are better able to check for and reduce errors that individuals might make. Vollrath et al. (1989), in a mock trial scenario, found that groups tend to perform a memory task better than individuals. Their study suggested that groups may reduce individual errors of commission and omission but exaggerate errors of implication. However, Clark and Wade (1986) found that groups were more likely to correct implicational errors than individuals. These conflicting results are typical of the theories that suggest groups are able to check for and reduce the individual errors. It should be noted that the current study will not specifically address this issue since there is no way of

determining if an improvement in group performance was due to a reduction of errors.

Pooling of information. A more promising theory that attempts to explain why groups typically perform better than individuals is that they are able to pool their informational resources. This type of theory has generally been supported by experimental research. Vollrath et al. (1989), studied group decision making in a mock trial scenario, and found that the greater the variability between members' estimates the more the group improved over homogeneous groups. Variability across subjects leads to more total information available. Hinsz (1990) found that group performance (d'_{grp}), while evaluating a job interview, was much better than individual performance (d'_i). Both of these findings support the theory that groups benefit from the pooling of independent information.

The greatest improvement in performance is obtained by the optimum combination of independent informational sources. The Theory of Signal Detection provides a measure of how well one may discriminate whether a sample was drawn from one of two normal distributions. The d' statistic represents this value

$$d' = \frac{\mu_s - \mu_n}{\sigma} \tag{1}$$

Where μ_s is the mean of the signal distribution, $\mu_n = 0$, and σ is the standard deviation for both distributions. Suppose the samples are drawn n independent times and are

added together, the mean of the summed distribution is $n * \mu_s$ and the variance is $n * \sigma^2$. Then

$$d'_n = \frac{n * \mu_s}{\sqrt{n} * \sigma} = \sqrt{n}\, d'_i \tag{2}$$

This is the maximum gain in performance available due to independent, normal informational sources. Partially correlated informational sources decrease both the variability (total information) and the potential increase in group performance. The pooling of information theory states that the more abilities and memories that are brought together on a task, the better the group performance.

Correlation between sources and observers. Another question related to adding group members, is how correlation between information sources (Sniezeck, 1990), as well as the correlation between the group members (Sniezeck & Henry, 1989), affects group performance.

An aviation example may be helpful to explain how information sources can be correlated. If two airplane engine gauges display the exact same information, by definition, they are perfectly correlated. The addition of the second gauge results in a zero gain in available information and therefore a crew cannot improve their performance. The other extreme, when group members receive independent information sources, provides the maximum amount of information available to the group.

Sniezeck (1990) found the performance of a group of experienced weather forecasters, presented with highly

correlated information, to be no better than individual performance. In a related finding, Johnson and Torcivia (1967) reported that a high correlation of knowledge related to performance within a group resulted in reduced group performance. They theorized this occurred because each individual contributed mainly common and little unique information. Even though the correlation between information sources has been demonstrated to be an important determinant in group performance, it is often not addressed in group research (for example, see Hinsz, 1990). If members of a group are from similar social and economic backgrounds it is likely that their performance and responses will be correlated and thereby reduce the variance in estimates across the group. This reduction in variance reduces the possible group performance. Therefore, group members from different backgrounds may increase the group performance more than group members with similar backgrounds.

Hogarth (1978), in a study of group performance, made an interesting suggestion. He indicated that, in order to raise group performance, it may be beneficial to add an observer who lowers the correlation between the group members even if that new member is not the best observer. The correlation between group members will be measured in the current study and its specific effect will be discussed.

Relative Weights

During the group decision process, we assume that the group makes a decision by differentially weighting the input

of each of the individual members. Measuring the ways that groups distribute relative weights (emphasis) to individual members during the group decision process is an important topic in group research. Optimal group performance is accomplished by assigning relative weights to the group members based on their individual performance (Grofman, Owen & Feld, 1983; Green & Swets, 1966; Green, 1994). This statistical calculation states that relative weights should be assigned to the group members monotonically with the individual's performance. By measuring deviations from ideal weights we can measure one way in which groups differ from the ideal.

Sniezeck and Henry's (1989) study of individual versus group judgments about frequency of death for different causes provides support for the theory that groups performing a group judgment do more than average the individual inputs or take the median response. They found that a model assuming equal weighting across group members provided a poor fit of their data. Their findings suggest that the groups weighted individual members input based on some criteria, possibly individual performance.

Individual subject's weighting. While the study of weights assigned to individual members during a group decision has been limited, studies of weighting strategies used by an individual are more frequent. These studies have measured the relative weights that individuals assign to different information sources during experimental trials in

order to make a decision. For instance, the results of Montgomery and Sorkin (in press) demonstrated that individuals, given cues, can weight visual sources based on their difficulty. In addition, Berg and Green (1991) showed that listeners can concentrate on the most important spectral channel in a hearing detection task. Furthermore, Sorkin et al. (1991) concluded that, given enough time, individuals can appropriately weight multiple visual elements.

Prior group research has made several important findings, and significant strides in understanding the group decision process continue to be made (Gonzales, 1994; Samuelson & Allison, 1994). Both of these studies looked at the influence that group members' actions had on one another. Since there are still several interesting questions about the group decision weighting process, the current study will measure the groups' ability to weight observers based on their individual performance abilities and the difficulty of their respective observations.

Ideal groups should weight the estimates of the group members according to each individuals' performance and the variability of the information the individuals observe. Returning to the aviation example, when both members of a crew receive information that is similar in difficulty, to discriminate signal from noise, the instructor pilot (who has shown excellent performance) should be given more relative weight than the new student pilot. However, if the instructor is reporting information from a gauge that is

difficult to discriminate (displays with large variability), his input should receive less weight than a pilot who is reporting an observation from a gauge that is less difficult to discriminate.

Comparison of Human and Ideal Performance

Prior studies of performance in display based visual signal detection tasks (Sorkin, Mabry, Weldon, & Elvers, 1991; Montgomery & Sorkin, in press; Ashby & Maddux, 1992) demonstrated that individual observers, after extensive training, can become accurate and consistent decision makers. Montgomery and Sorkin (in press) found that individual observers performing a discrimination task, using a multi-element visual display, were able to effectively differentiate and weight independent gauge values based on their variability. The subjects correctly assigned more relative weight to the display values that had smaller variance than to values with higher variance. Furthermore, individual performance (d'), on the discrimination task depended on display variability in a predictable way. This research suggests that individuals can perform visual detection tasks effectively; however, it is still uncertain as to how efficiently groups can perform these tasks.

Sorkin and Dai (1994) used the Theory of Signal Detection to calculate the normative equations that represent ideal group performance. The following section summarizes their analysis of the Ideal Group. All assumptions and equations listed are from their analysis unless stated

otherwise. The current study will use their analysis to evaluate how group performance compares to an ideal group.

The assumed group decision process is shown in Figure 1. The members in a group are all presented information, and they have to decide if the stimuli were representative of a noise or signal trial. The stimuli presented to the group members, on a given trial, are drawn from the same probability distributions, but are not necessarily identical across subjects. During the trial, each observer may be subject to several noise sources. These noise sources are divided into two categories: those that are unique to each observer (in Figure 1 n_1, n_2, n_3) and those noise sources that are common to two or more observers ($n_{1,2,3}$, $n_{1,3}$). The noise sources are all assumed to be additive, normally distributed Gaussian random variables with zero mean and variances of $\sigma_1{}^2$, $\sigma_2{}^2$, $\sigma_3{}^2$, $\sigma_{1,2,3}{}^2$, $\sigma_{1,3}{}^2$.

After the stimulus presentation each observer offers her/his estimate, x_i, to the group. These estimates are also assumed to be normally distributed Gaussian random variables with a mean of zero on noise trials and a mean of μ_i on signal trials. These estimates from the individual observers $<x_1, x_2, x_3>$ form the group estimate vector, $\mathbf{x}$.

Sorkin and Dai show that the weighted sum of the group members' estimates (x_i) forms an optimal (likelihood ratio) decision statistic, $Z = \Sigma a_i x_i$. The group then compares this decision statistic to a pre-set criterion, Z_C, to decide if they should answer signal or noise. The group decision

problem is analogous to that of detecting a signal that has components in m channels. This decision problem is similar to the multi-channel auditory signal detection problem examined by several other researchers (e.g. Berg, 1989, 1990; Berg & Green, 1990; Durlach et al., 1986; & Green, 1988 , 1992).

The detection index of each group member, d'_i, is a measure of how well the subject can detect signal from noise. For example, on Figure 1,

$$d'_1 = \frac{\mu_1}{\sqrt{\sigma^2_1 + \sigma^2_{1,2,3} + \sigma^2_{1,2}}} \tag{3}$$

The relative weights assigned to each group member during the group decision process is designated as a_i. The correlation between the group members' estimate is called the inter-member correlation.

A simplifying assumption is that all the unique variances are equal in magnitude across the members

$$\sigma^2_i = \sigma^2_{ind} \text{ for all } i. \tag{4}$$

The correlation between two members is the ratio of the common (shared) variance to the total variance

$$r = \sigma^2_{com} / (\sigma^2_{ind} + \sigma^2_{com}). \tag{5}$$

Each group members' total variance is normalized by setting $\sigma^2_{ind} + \sigma^2_{com} = 1$. Then, from the definition of correlation; $\sigma^2_{com} = r$ and $\sigma^2_{ind} = 1 - r$. Using equations derived by Durlach, Braida and Ito (1986), Sorkin and Dai (1994) showed that the performance of the ideal group is:

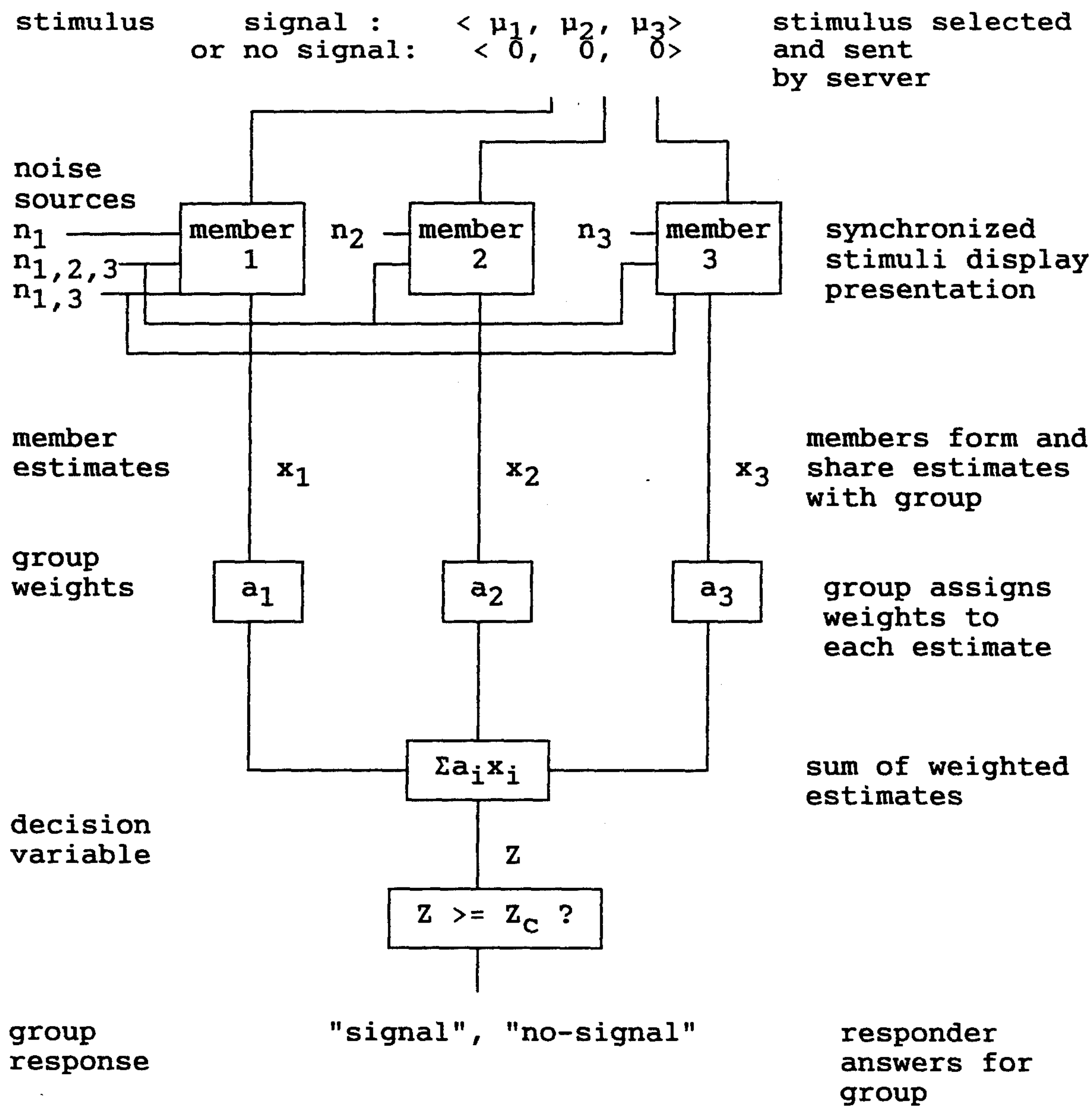


Figure 1. Assumed group decision process (after Sorkin & Dai, 1994). Each group member experiences a unique noise source plus noise that is shared with one or more of the group members. The noise sources are assumed to be independent random normal variables with zero means and specified variances. The variances are independent of which stimulus event was presented. The group decision variable, Z, is the weighted sum of the group members' estimates.

$$d'_{\text{Ideal Group}} = \left[\frac{m\ \text{Var}\ (d')}{1 - r} + \frac{m\ (\bar{d}')^2}{1 + r\ (m-1)}\right]^{\frac{1}{2}}. \tag{6}$$

This important equation specifies ideal group performance, where m is the number of group members, $\bar{d}'$ is the mean of all d'_i, and Var (d') is the variance of the d'_i. In the case where the group members' estimates are independent $(r = 0)$, then

$$d'_{\text{Ideal Group}} = [m\ \text{Var}(d') + m\ (\bar{d}')^2]^{\frac{1}{2}} \tag{7}$$

$$= [\sum_{i=1}^{m} (d'_i)^2]^{\frac{1}{2}}. \tag{8}$$

This is the independent channel prediction made by Green and Swets (1966) and is equal to $\sqrt{m}\, d'_i$ when all d_i are equal.

The Ideal Group analysis uses four primary variables to specify group performance. These variables are: number of group members (m), inter-member correlation (r), and the mean and variance of the group members' d's. These variables can be thought of in the aviation example. The d'_i of the group involves how well the crew members can determine if they are experiencing problems. The difficulty of discriminating signal from noise and the attentiveness and skill of the pilots is involved in this measure. The correlation between the group members' estimates is given by r. The variable r involves the training and backgrounds of the crew members and whether the crew members are monitoring common or independent sources of information.

By looking at the equation describing ideal performance, we can see the effects of each variable. As d_i, m, and Var d' increases so does $d'_{Ideal\ Group}$. However, when inter-member correlation (r) increases, $d'_{Ideal\ Group}$ decreases.

Sorkin and Dai (1994) show the ideal weights ($\hat{a}_i$) to be

$$\hat{a}_i = \frac{1 + r\,(m-1)}{r}\, d'_i - m\,\bar{d}' \tag{9}$$

The ideal weight for each group member is directly proportional to the individual's d'_i.

Often it is informative to determine the relative efficiency of an actual group. The efficiency of a group (from Tanner & Birdsall, 1958) is defined as the ratio of the squared d' of the group to the squared d' of the Ideal Group.

$$\eta_{observed} = \frac{(d'_{actual\ group})^2}{(d'_{Ideal\ Group})^2} \tag{10}$$

The reason for using this definition, is that in many psychophysical experiments, the ideal $(d')^2$ is proportional to signal energy. Therefore, an efficiency level of .60 means that the Ideal Group would need only 60% of the stimulus energy to perform at the same level as the actual group. Ideal performance would result in efficiency = 1.0 and any performance less than ideal would result in $\eta < 1.00$.

Observed efficiency can be separated into two efficiency measures to determine why the group suffered losses in group performance. First, weighting efficiency is defined as d'_{weight} squared over $d'_{Ideal\ Group}$ squared. The measure d'_{weight} assumes ideal performance but uses the actual weights used by the group (Berg,1990).

$$\eta_{weight} = \frac{(d'_{weight})^2}{(d'_{Ideal\ Group})^2} \tag{11}$$

Weighting efficiency measures the loss in group performance (d'_{grp}) due to less than ideal weights. By manipulating the equations given in Sorkin and Dai's paper, d'_{weight} for the actual group is:

$$d'_{weight} = \frac{\Sigma a_i\ d'_i}{\sqrt{(1-r)\Sigma a_i{}^2 + r(\Sigma a_i)^2}} \tag{12}$$

Or in the uncorrelated case, the same analysis shows that

$$d'_{weight} = \frac{\Sigma a_i d'_i}{\sqrt{\Sigma a_i{}^2}} \tag{13}$$

The second efficiency measure (Berg, 1990), η_{noise}, accounts for additional loss in d'_{grp} due to reasons besides weights,

$$\eta_{noise} = \frac{(d'_{observed\ group})^2}{(d'_{weight})^2} \tag{14}$$

The relationship between the efficiency measures (Berg, 1990) is

$$\eta_{obs} = \eta_{weight} * \eta_{noise}. \tag{15}$$

Summary

The effect of group size has been studied by several researchers (Grofman, Owen & Feld, 1983; Sniezeck & Henry, 1989; Vollrath, Sheppard, Hinsz & Davis, 1989). Group performance is generally found to be superior to that of an individual (Hartwick et al, 1982). Specific findings are: performance improves with the addition of observers (Johnson & Torcivia, 1967); it is better than simply averaging the individual responses (Sniezeck & Henry, 1989); and that it exceeds the performance of the best member (Sniezeck & Henry, 1989). However, the specific increase in performance gained by adding an additional member has not been accurately predicted (Vollrath et al, 1989).

Several theories have been used to speculate as to why groups typically perform superior to individuals. The Best Member and Error Checking theories have yielded mainly unfavorable results, whereas the Pooling of Information theory seems to provide a more likely explanation.

Research from several fields ranging from jury research (Marquis de Condorcet, 1785) to group memory research (Hinsz, 1990) to psychophysics (Montgomery, 1993; Berg & Green, 1991, Sorkin et al., 1991) suggests that a more complete method of analysis for group decision research is necessary. The current study will investigate previously studied variables that have shown to affect group performance. Also, actual group performance will be compared against an ideal to determine how efficiently groups per-

form. It is hoped that these comparisons may be used to benefit actual decision-making groups. By demonstrating typical losses in possible group performance, decision-making groups can then be trained to be aware of and to avoid them.

METHOD

The subjects performed, either individually or as a group, a visual signal detection task. The subjects were presented a multiple element visual display screen, composed of nine analog gauges (shown in Figure 2), for a short period of time (320 ms) and asked to respond signal or noise. The subjects had to discriminate if the stimulus display screen was representative of samples drawn from a normal distribution with mean = μ_s (signal) or with mean = μ_n (noise). Both of the distributions had a standard deviation = σ. The means of the two distributions are shown as double tick-marks in Figure 2 ($\mu_s > \mu_n$). In the current study, μ_s = 5.0 and μ_n = 4.0.

The difficulty of the trials was controlled by adjusting the standard deviation of the noise and signal stimulus distributions, σ. A stimulus display screen that exhibited large variability, for example σ = 3.0 (each gauge), would make discriminating signal from noise difficult. Using a stimulus display screen with small variability, for example σ = 1.5, would make it easier to discriminate which distribution was used to generate the stimulus display.

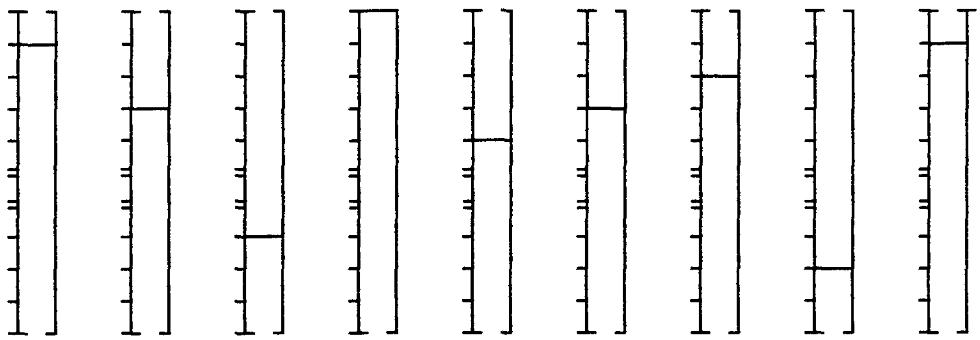


Figure 2. An example of the nine vertical gauges presented on a stimulus display screen. This screen was presented to an individual during an individual subject trial or as part of a group during a group trial.

Subjects

Eight University of Florida students, seven women and one man, 19 to 23 years of age participated in this study. The subjects were paid 4.25 dollars per hour. A small incentive bonus was paid based on the subjects' performance (approximately 40 cents per hour). All of the subjects had normal or corrected-to-normal visual acuity. The subjects were tested both individually and in groups of different sizes.

Apparatus and Stimuli

Insight 486-33dx computers were used to generate and present the stimulus display screens via a 14 inch CTX color monitor (1024 X 768 SVGA monitor, 72 Hz refresh rate at 640 X 480 resolution). Observers sat approximately 27 inches away from the monitor in a semi-quiet, fluorescent lighted laboratory room. The monitor was set for maximum contrast, and the intensity was set at approximately 100 cd/m^2, measured from a 7.5 inch by 10.5 inch uniform white field. During each trial, nine vertical gauges were presented subtending a visual angle of approximately 8° by 16° (vertical by horizontal). Responses were made via a standard computer keyboard.

The individual gauges, shown in Figure 2, consisted of two white, parallel, vertical lines with tick marks on the left line dividing the gauge equally in one number increments from 0.0 to 10.0. Two larger blue tick marks marked the mean of the noise (μ_n) and signal (μ_s) distributions

(4.0 and 5.0 respectively). During a given trial, all of the gauges on a stimulus display screen displayed values representative of the same distribution. The stimuli were randomly selected from the signal distribution on half of the trials and from the noise distribution on the other half of the trials. These values were then presented as horizontal, white dashes on the gauges, where 0.0 represented the bottom line of the gauge and 10.0 represented the top line of the gauge.

The duration of the stimulus presentation was synchronized with the refresh traces of the monitor. The onset and offset of the stimulus presentation were delayed until it was time for another screen refresh (approximately every 13 ms). The stimulus duration was 370 ms for two practice sessions. The remainder of the trials had a stimulus duration of 320 ms.

During the group phase of the experiment, the subjects sat in front of their individual monitors; the group members' stimulus display screens were driven simultaneously via a Local Area Network (LAN). This prevented subjects from viewing screens other than their own. A 486-33dx computer acted as the "mini-network" server and controlled the experiment by synchronizing the trials and recording the group's performance and conditions. During each trial, this computer randomly assigned one of the group members to respond for the group.

All of the group interaction trials, except the eight member group, took place in the same room as the individual trials in an L-shaped configuration. The server computer was approximately five feet from the closest member's monitor along one leg of the L. Seating arrangements during interaction for the group of four were randomized across sessions with two members side by side (separated by 24 inches), sitting perpendicular to the other two members (also separated by 24 inches), and separated at the corner by approximately 30 inches. In the groups of two, the members sat facing the same direction with their monitors separated by 24 inches. These seating arrangements were chosen to promote open interaction and easy communication between the group members.

The large group trials took place during the evening, with similar lighting and sound levels as the experimental room. During the large group trials, half of the members sat side by side, along one wall of the room, while the other half of the members sat on the other side of the room (approximately ten feet away). The members sitting side by side were separated by small wooden partitions and were approximately 24 inches apart. The two subsets of the group had their backs towards each other with the server computer on one side of the room. Again, this seating arrangement was chosen over other possible arrangements to promote discussion and make for the best achievable interaction.

Procedure

Each subject was first tested individually, in order to determine the baseline performance on the task. After the individual testing, subjects were divided into groups of various sizes to perform the group task. These participants were randomly divided to form one group of eight members, two groups of four members and two groups of two members. The two member groups were randomly chosen from the four member groups. However, the male subject was purposely excluded from the group of two observers to minimize the chance that his presence would bias that group's decision (Clement & Schiereck, 1973). The large group consisted either of six or seven members due to absenteeism and scheduling problems. All subjects were retested in the individual task at the end of the study.

Individual Subject Sessions

The subjects were instructed to respond signal or noise on each trial depending on the height of the stimuli on the nine independent gauges relative to the noise and signal mean tick-marks. The subjects were asked to press either the "2" key for noise or the "3" key for signal, for the representative "noise" and "signal" stimulus display screen.

First, each subject's individual baseline performance was determined through several individual sessions. A representative trial is shown in Figure 3. First, a 0.5" by 0.5" white fixation cross was presented in the center of the display for 200 ms. Next, the nine visual display

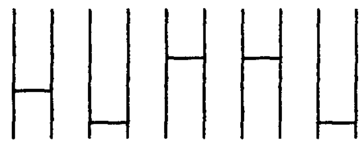
Trial Sequence	Display Screen
Fixation: 200 ms	+
Stimulus Duration: 370 ms (initial practice) 320 ms (test)	
Masking Screen: 200 ms	White Screen
Response Duration: 1000 ms	Blank Screen
Feedback: 250 ms	Incorrect

Figure 3. Trial sequence for individual subject testing.

gauges were presented for 320 ms. Immediately after the stimulus display screen, a 200 ms white masking screen was presented. Following the mask, the screen was completely blank for one second at which time the subject could press "2" for noise or "3" for signal. If an answer was input before or after the allotted one second response period, the response was discarded and "No Response" feedback was given. Finally, the observer received feedback in the center of the screen for 250 ms. The given feedback was either "Correct", "Incorrect", or "No Response".

Each individual subject session consisted of 16 blocks of 125 trials per block for a total of 2000 trials per session. Four difficulty conditions were utilized in the current study: $\sigma = 3.0$, $\sigma = 2.5$, $\sigma = 2.0$, and $\sigma = 1.5$. Across four sessions, each subject completed 2000 trials (1000 signal and 1000 noise) for each of the four difficulty conditions. All subjects received two practice sessions with a 370 ms stimulus duration. In addition, seven subjects received an additional two to four practice sessions with a 320 ms stimulus duration. However, subject S_7 received only the initial practice sessions. Furthermore, S_7 was only able to complete three sessions for a total of 1500 trials for each difficulty condition while the other subjects completed 2000 trials. All the subjects were highly practiced at the discrimination task prior to recording individual data.

To control for possible practice effects, the four difficulty conditions were randomly assigned in four block combinations. Each subject completed four blocks of each difficulty during a session. Across sessions, the four block sequences were also randomized. For instance, a session could occur in the following order; Four blocks of trials with $\sigma = 1.5$, four blocks of trials with $\sigma = 2.5$, followed by four blocks with $\sigma = 2.0$, and finally, four blocks with $\sigma = 3.0$.

The subjects were instructed to take breaks after each block in order to stay attentive. Each experimental session took approximately 1.5 hours to complete.

Group Sessions

After the individuals' data were collected, the members were tested in groups of two, four, and eight members. The experimenter instructed the groups to make a single group response for each of the trials. In order to allow for free interaction among group members, the method of interaction and decision rule was not specified and was left to the separate groups.

During a group trial, each subject was presented with a stimulus display screen, representative of either signal or noise. Figure 4 illustrates an example of a group trial. The stimulus displays presented to the groups on group trials were the same as those presented on individual trials with the following exceptions. After the white masking screen, all the group members' computer screens were com-

pletely blank for 700 ms. This blank screen period was intended to signal the start of group discussion about the trial. Following the 700 ms pause, one of the group members was randomly selected and signaled on her/his computer screen to respond for the group. Unlike the individual trials, no time limit was imposed for the group discussion and interaction. Therefore, none of the trials were discarded as "No Response" trials. After the group response, a 1000 Hz tone with a duration of 150 ms was generated by the server computer. The purpose of this tone was to inform the members that a group response had been made and to focus their attention on their monitors. At that time, the feedback was presented for 600 ms. Feedback on the group trials included the group response in addition to "Correct" or "Incorrect". The experimenter either left the room or sat quietly away from the subjects to take notes about the group interaction. The subjects could not see what the experimenter was working on during their interaction.

Inter-member correlation. As in the individual conditions, the stimuli presented on a single stimulus display screen (for a single group member) were always made up of nine independent, normal random variables. The stimuli presented on the stimulus display screens (one for each group member) were generated by a process that allowed the correlation between a pair of group member's stimuli to be controlled. The stimuli were either independent ($r = 0.0$) or correlated ($r = 0.25$) across group members.

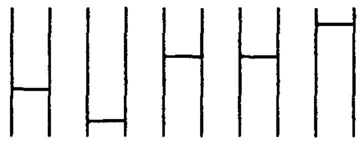
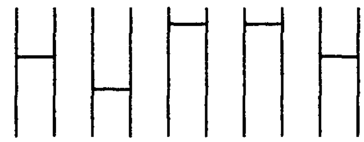
Trial Sequence	Responder	Non-Responder
Fixation: 200 ms	+	+
Stimulus Duration: 320 ms		
Masking Screen: 200 ms	White screen	White screen
Start of Interaction: 700 ms	Blank screen	Blank screen
Group Responder Randomly Chosen: after 700 ms of blank screen	You have been selected for a response station. Please answer the following question Did you think that was signal or noise?	Blank screen
No time limit for response		
Server sounds tone after response		
Feedback: 600 ms	Response: Noise Incorrect	Response: Noise Incorrect

Figure 4. An example of the synchronized trial sequence for group testing (two-member group shown).

The statistical process that correlated the stimulus values is shown in Figure 5. As in Sorkin and Montgomery (1991), the stimulus values for a spatially positioned gauge (1..9) for two group members were produced by combining three independent, normal random variables X_a, X_b, and X_c where $\mu_a = \mu_b$, and $\sigma^2{}_a = \sigma^2{}_b \equiv \sigma^2{}_u \equiv \sigma^2{}_c$. Where $\sigma^2{}_u$ is defined as the unique variance and $\sigma^2{}_c$ is defined as the common or shared variance. To control the correlation between the stimuli values presented on the first gauge position for two group members, we form the two values $(G_{1,i})$ and $(G_{2,i})$. These values are produced by taking samples of the random variables G_1 and G_2. $G_{1,i}$ and $G_{2,i}$ are defined according to the three independent, normal random variables; X_a, X_b, and X_c. Let

$$G_{1,i} = uX_a + cX_c \tag{16}$$

$$G_{2,i} = uX_b + cX_c. \tag{17}$$

Where

$$E(G_{1,i}) = E(G_{2,i}) = \mu_c$$

and

$$\mathrm{var}(G_{1,i}) = \mathrm{var}(cX_c) + \mathrm{var}(uX_a) \tag{18}$$

$$= c^2\sigma^2{}_c + u^2\sigma^2{}_a = (c^2 + u^2)\,\sigma^2{}_c \tag{19}$$

$$\mathrm{var}(G_{2,i}) = \mathrm{var}(cX_c) + \mathrm{var}(uX_a)$$

$$= c^2\sigma^2{}_c + u^2\sigma^2{}_b = (c^2 + u^2)\,\sigma^2{}_c.$$

Then, the correlation, r, between $G_{1,i}$ and $G_{2,i}$ is

$$r_{G1,i,G2,i} = [\mathrm{cov}(G_{1,i},G_{2,i})]/\sigma_{G1}\sigma_{G2}, \tag{20}$$

$$\mathrm{cov}(G_{1,i},G_{2,i}) = E[(G_1-\mu_1)(G_2-\mu_2)] \tag{21}$$

$$= E[(G_{1,i}G_{2,i} - \mu_1 G_{2,i} - \mu_2 G_{2,i} + \mu_1\mu_2)]$$

assuming noise trials where μ_1 and $\mu_2 = 0$,

$$= E(cX_c + u X_a)(cX_c + uX_b)$$

$$= E[c^2X_c^2 + u^2X_aX_b + uX_acX_c + uX_bcX_c]$$

since the distributions are independent

$$= E\ [c^2X_c^2] = c^2\sigma_c^2 \qquad (22)$$

then

$$r_{G1,G2} = \frac{c^2\sigma^2{}_c}{\sqrt{(c^2 + u^2)\sigma_c^2}\ \sqrt{(c^2 + u^2)\sigma_c^2}} = \frac{c^2}{c^2 + u^2}. \qquad (23)$$

It is convenient to let

$$c^2 + u^2 = 1, \qquad (24)$$

then

$$c = \sqrt{r} \text{ and} \qquad (25)$$

$$u = \sqrt{1 - c^2}. \qquad (26)$$

In Figure 5 the stimulus values displayed on gauge #1 for two members are made up of a sample of a shared random variable (X_c) and samples of unique random variables (X_a and X_b). From the equations above, stimulus values that share a large portion of the variance ($\sigma^2{}_c$) are more likely to have similar display values than the stimulus values that have no common variance. In order to correlate all gauge positions, the same process was repeated for each gauge (1..9). Additionally, all possible pairs of members' stimulus values were generated in the same manner so that all the inter-member correlations could be controlled.

The statistical process correlated each like spatially positioned gauge between group members during the r = 0.25 condition. During the r = 0.0 condition, the stimulus

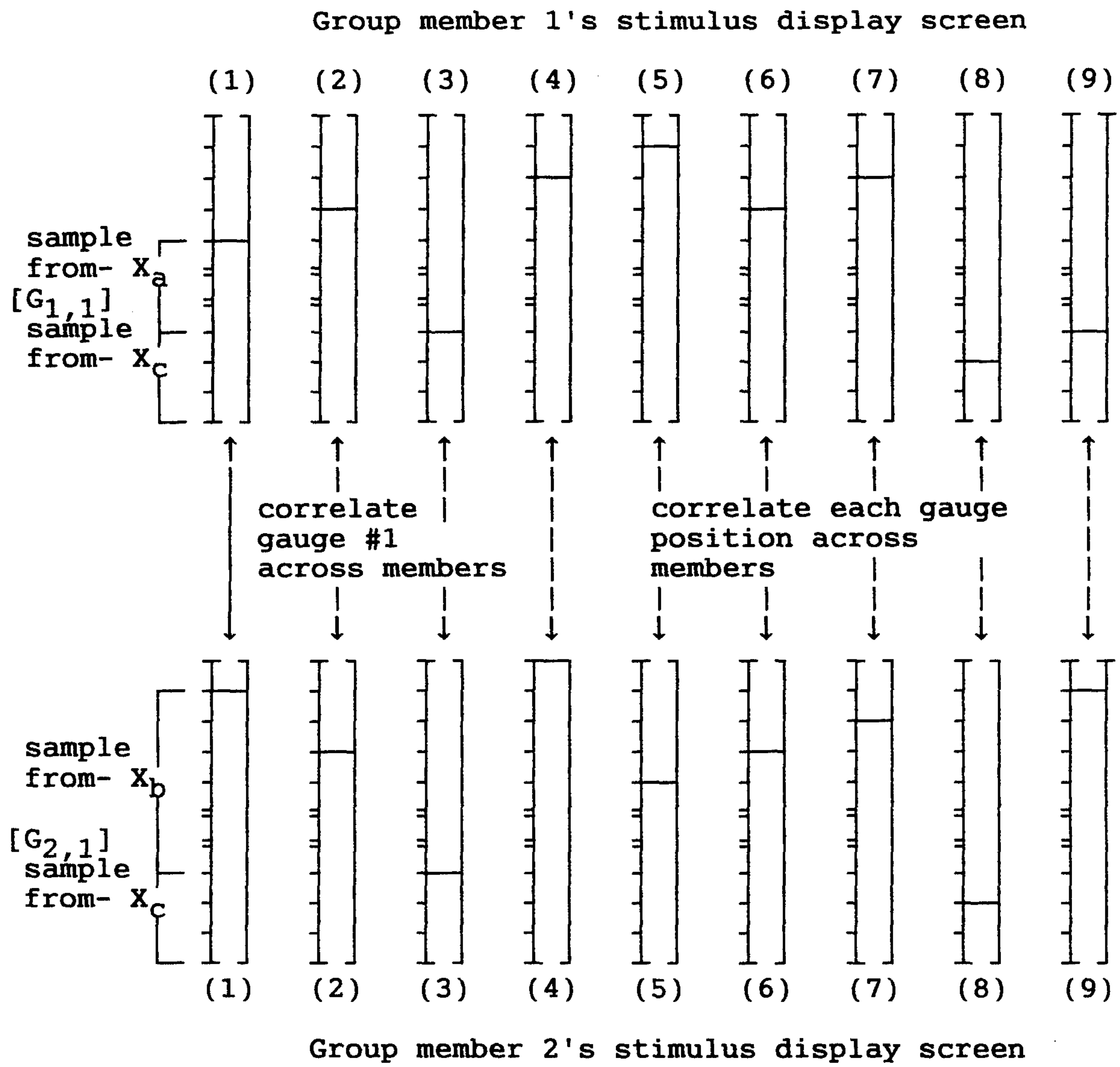


Figure 5. Graphical representation of how stimulus values presented on the same gauges were correlated across group members during the r = 0.25 condition (trial #1). Gauge locations are labeled in parenthesis.

values between group members were made up entirely of samples of the unique random variables. In this manner, the amount of shared variance was experimentally controlled. Since each gauge position across each pair of group members was correlated, the entire group members' observations were statistically correlated. The inter-member stimulus correlation was verified with a Monte Carlo simulation.

Experimental conditions for group trials. Each experimental session, for the group trials, consisted of eight blocks of 100 trials. The group session was separated into two, four-block sets, with the same condition held constant during a four-block set. In order to account for possible practice effects, the first block of trials for each new group condition was labeled practice and not evaluated. During the equal difficulty distribution trials, all group members observed stimulus display screens with the same difficulty (σ) condition. For group trials, the specifications of each condition consisted of an inter-member correlation (r) and the difficulty level (σ) for each group member. Table 1 depicts a listing of all the trial conditions. The stimuli between members were either correlated ($r = 0.25$) or independent ($r = 0.0$). The distribution of difficulty of the trials across group members was either equal or unequal and is listed in the < >. Both the difficulty and correlation aspects of the trial condition were randomized between sets and across sessions.

Table 1.
Experimental conditions for group trials.

Group Size	Inter-member Correlation (r)	Difficulty Distribution <σ>
2	0.00	Equal σ : <1.5,1.5> or <3.0,3.0>
	0.25	
4	0.00	Equal σ : <2.0,2.0,2.0,2.0> or <3.0,3.0,3.0,3.0>
		Unequal σ : <1.5,1.5,3.0,3.0>
	0.25	Equal σ : <2.0,2.0,2.0,2.0> or <3.0,3.0,3.0,3.0>
6	0.00	Equal σ : <2.5,2.5,2.5,2.5,2.5,2.5> or <3.0,3.0,3.0,3.0,3.0,3.0>
	0.25	
7	0.00	Equal σ : <2.5,2.5,2.5,2.5,2.5,2.5,2.5> or <3.0,3.0,3.0,3.0,3.0,3.0,3.0>
	0.25	

The number of replications and membership for the different groups are listed in Table 2. The number of 100-trial blocks is listed in parenthesis after the group membership. The two and four member groups completed at least two, four-block sets (400 signal trials and 400 noise trials) for each of the selected trial conditions. The conditions during the large group interaction were not replicated with the same exact members in the group each time. Depending on the size of the group, some of the difficulty conditions were omitted because the group performance would have been too high to measure accurately. The groups were instructed to take breaks in between blocks to stay alert and attentive. The group sessions, similar to the individual sessions, took approximately 1.5 hours to complete.

Unequal difficulty condition. In addition to the equal difficulty distribution trials, a single group of four subjects completed four, four-block sessions where the distribution of difficulty (σ) of trials, across the subjects, was unequal. During a set of four blocks, two members observed stimulus display screens with small variability ($\sigma = 1.5$) while the other two members observed stimulus display screens with larger variability ($\sigma = 3.0$). The different difficulty levels of the task were randomly assigned based on the first two subjects logged in and were held constant for each four-block set. As in the equal difficulty trials, an unequal difficulty session consisted

Table 2.
Group membership and number of replications for group trials in all experimental conditions. Number of blocks for each condition are listed inside of parenthesis. There were 100 trials per block. All groups were under equal difficulty (σ) distribution except the four member group.

2A. 2 Member Groups

Difficulty	Inter-member correlation (r)	
σ	0.0	0.25
<1.5,1.5>	S_1 S_2 (12) S_5 S_6 (8)	S_1 S_2 (8) S_5 S_6 (8)
<3.0,3.0>	S_1 S_2 (8) S_5 S_6 (8)	S_1 S_2 (8) S_5 S_6 (8)

2B. 4 Member Groups

	Difficulty	Inter-member correlation (r)	
	σ	0.0	0.25
EQUAL	<2.0,2.0,2.0,2.0>	S_1 S_2 S_3 S_4 (8) S_5 S_6 S_7 S_8 (8)	S_1 S_2 S_3 S_4 (8) S_5 S_6 S_7 S_8 (8)
	<3.0,3.0,3.0,3.0>	S_1 S_2 S_3 S_4 (8) S_5 S_6 S_7 S_8 (12)	S_1 S_2 S_3 S_4 (8) S_5 S_6 S_7 S_8 (11)
UNEQUAL	<1.5,1.5,3.0,3.0>	S_7 S_8 S_2 S_5 (4) S_2 S_5 S_7 S_8 (8) S_5 S_7 S_2 S_8 (4)	

Table 2 -- continued

2C. <u>6 and 7 Member Groups</u>

Difficulty	Inter-member correlation (r)	
σ	0.0	0.25
<2.5,..,2.5>	S_1 S_3 S_5 S_6 S_7 S_8 S_1 S_2 S_3 S_5 S_6 S_7 S_8 S_1 S_2 S_3 S_4 S_5 S_6 S_7	S_2 S_3 S_4 S_5 S_6 S_7 S_8 S_2 S_3 S_4 S_6 S_7 S_8
<3.0,..,3.0>	S_1 S_3 S_5 S_6 S_7 S_8 S_2 S_3 S_4 S_5 S_6 S_7 S_8 S_2 S_3 S_4 S_6 S_7 S_8	S_1 S_2 S_3 S_5 S_6 S_7 S_8 S_1 S_2 S_3 S_4 S_5 S_6 S_7

All conditions listed for 6 and 7 member groups are based on four blocks of trials.

of a four-block set for each of two separate conditions. Table 3 lists the six possible unequal difficulty-subject combinations. The subjects completed a four block set for three of the six possible subject-difficulty combinations with combination #2 being replicated.

All unequal difficulty distribution trials had an inter-member correlation (r) = 0.0. The same instructions were given to the unequal σ group as had been given to prior groups. The subjects were not told about the differences in difficulty across subjects. All of the equal difficulty trials had been completed prior to the start of the unequal difficulty distribution group.

<u>Individual Retest</u>

After the equal difficulty group trials, the subjects were retested in the four difficulty conditions. The goal of this retest was to determine whether changes in individual performance (d'_i) occurred during the period of the group trials. The individuals were tested under each difficulty level for four blocks of 125 trials per block. The procedure and instructions for the retest were the same as the original individual trials.

Table 3.
The six possible unequal difficulty-subject test combinations. The tested conditions are within the box. Combination #2 was replicated twice for a total of 8 blocks. Combinations #1 and #3 represent 4 blocks of trials.

		SUBJECT NUMBER			
		S_2	S_5	S_7	S_8
	#1	3.0	3.0	1.5	1.5
	#2	1.5	1.5	3.0	3.0
DIFFICULTY COMBINATIONS <σ, σ, σ, σ>	#3	3.0	1.5	1.5	3.0
	#4	1.5	3.0	1.5	3.0
	#5	1.5	3.0	3.0	1.5
	#6	3.0	1.5	3.0	1.5

RESULTS

This section begins with consideration of the individual performance (d'_i) measured before the group trials and during a subsequent retest. Next, the groups' performance (d'_{grp}) is analyzed based on task difficulty (σ), the number of members in the group (m), and inter-member correlation (r). Consideration of these three effects is followed by the groups' weighting strategies along with some possible weighting problems. Finally, an analysis of the groups' observed ($^n{}_{obs}$) and weighting ($^n{}_{weight}$) efficiencies is presented.

Individual Results

During the individual trial sessions, the individual's responses were recorded, along with the presented stimuli (signal or noise). These values were used to calculate hit and false alarm probabilities for a block of 125 trials. These probabilities were then transformed to a measure of the individual's detection performance, d'.

The mean individual performances (d'_i), across the four difficulty levels, are listed in Table 4. Performance differences between subjects were not significant ($F(4,8)=.348$ $p<.923$). This can be seen in Table 4, as performance differences due to the difficulty conditions were greater than most of the differences between individual subjects.

Subject performance (d') decreased with increases in difficulty (σ) ($F(4,8)=61.56$ $p<.001$). A post-hoc analysis (Tukey's Test) confirmed that all the difficulty conditions were significantly different (d' decreased with each increase in σ, $p<.05$). There was no significant interaction between difficulty level (σ) and subject.

The ideal performance for an individual subject, d'_{ideal}, within each of the difficulty conditions is listed in Table 5. Since the stimulus values presented on each of the nine gauges were independent random samples, equation 27 was used to determine d'_{ideal}.

$$d'_{ideal} = [\sum_{i=1}^{n} (\frac{\delta_i}{\sigma_i})^2]^{\frac{1}{2}} \qquad (27)$$

Where δ_i refers to the difference between the means of the two stimulus distributions ($\mu_s - \mu_n$); in this study it is a constant equal to one. Additionally, σ_i refers to the standard deviation of one display gauge. The d'_{ideal} values presented in Table 5 can be used as a reference to determine how actual performance (Table 4) compared to ideal.

The subjects performed at approximately 60% efficiency levels across all difficulty conditions. Average efficiencies for the subjects in each difficulty condition were determined by squaring the ratio ($d'_{obtained}/d'_{ideal}$) (Tanner & Birdsall, 1958). The average observed efficiency (η_{obs}) for each difficulty level was : 59% for σ=3.0, 61% for σ=2.5, 64% for σ=2.0, and 59% for σ=1.5. These measures show that the subjects performed at approximately the same

Table 4.

Mean performance (d'_i) for individual subjects, tested prior to group trials, under the four different difficulty conditions. There were 125 trials per block, with values based on 12 to 17 blocks. The average performance (d') and standard deviation within each difficulty condition, across subjects, is listed at the bottom of the table. The average observed efficiency (η_{obs}), across subjects, for each difficulty level is also listed.

Difficulty Level (σ_i) (decreasing →)

Subject	(3.0)	(2.5)	(2.0)	(1.5)
S_1	0.83	0.99	1.36	1.57
S_2	0.82	1.03	1.45	1.60
S_3	0.71	0.90	1.12	1.32
S_4	0.66	0.90	1.21	1.62
S_5	0.74	0.91	1.12	1.64
S_6	0.67	0.81	1.07	1.34
S_7	0.92	1.06	1.26	1.69
S_8	0.69	0.91	0.98	1.47
Average (d') Performance	0.76	0.94	1.20	1.53
Standard Deviation	0.092	0.082	0.156	0.139
Average (η_{obs}) Efficiency Across Subjects	0.59	0.61	0.64	0.59

Table 5.

The four difficulty conditions used during the individual and group trials. d'_{ideal} is given for the optimal performer during individual trials. During the individual trials, the sources of information (gauges) were always independent and therefore the optimum individual performance with n gauges is:

$$d'_{ideal} = \left[\sum_{i=1}^{n}\left(\frac{\delta_i}{\sigma_i}\right)^2\right]^{\frac{1}{2}} = \frac{\delta}{\sigma}(1)\sqrt{n}$$

Since the trials consisted of nine gauges and $\delta = (\mu_s - \mu_n)$ is a constant = 1, $d'_{ideal} = (3)(1/\sigma_i)$.

Difficulty Level (decreasing →)

Standard deviation σ_i	3.0	2.5	2.0	1.5
d'_{ideal}	1.0	1.2	1.5	2.0

efficiency level across the difficulty conditions. They also show that the subjects were able to effectively utilize the information on the gauges to perform the signal detection task.

Individual Retest

Following the equal σ group trials, the individual subjects were retested under the four different difficulty (σ) conditions, in order to determine if individual performance (d'_i) changed during the group trials. The retest session consisted of a four-block set for each of the four difficulty conditions, presented in random order.

The differences between mean individual performance prior to and after the group trials (measured in d') are shown in Table 6. Individual performance on the retest was approximately the same as measured on the pre-group trials, (mean cell change = 0.04). The retest showed that there was no significant change in individual d' values from the pre-group trials to the retest ($F(1,48)=.246$ $p<.622$). There was no significant difference between any of the subjects in the retest ($F(7,48)=.631$ $p<.728$), and the interaction between subjects and time of test ($F(7,48)=.224$ $p<.978$) was not significant.

Group Results

During each group trial, the sum of the stimulus gauge values presented on a display screen for each subject, the group response, and the stimulus (signal or noise), were recorded. The latter two values were then used to determine

Table 6.
Change in mean individual performance (d'_i) compared to original individual performance (Table 4) measured at the end of the group trials. There were 125 trials per block, with retest values based on 4 blocks. A decline in performance on the retest is indicated by (-).

Difficulty Level (σ_i) (decreasing →)

Subject	(3.0)	(2.5)	(2.0)	(1.5)	Mean Subject Change
S_1	.15	-.01	-.16	-.04	-.085
S_2	.22	.33	-.33	.21	+.173
S_3	.22	.18	-.06	.10	+.110
S_4	.20	.11	.03	.18	-.150
S_5	-.06	-.19	-.07	-.28	-.015
S_6	-.05	.14	-.04	.16	+.108
S_7	-.12	-.07	.03	-.18	+.053
S_8	.18	.08	.25	.18	+.130
Mean Difficulty Change	+.093	+.073	-.044	+.041	+.040 Mean Cell Change

hit and false alarm rates. The hit and false alarm rates were then transformed into detection performance for the group (d'_{grp}). The average group performance (d'_{grp}) for each trial condition is listed in Tables 7 and 8. Each group completed two, four-block sets during each experimental session. The experimental conditions (difficulty (σ) and inter-member correlation (r)) were presented in random order and held constant for each four-block set. The values for two and four-member groups are averages for two groups that completed at least 800 trials for each difficulty condition. However, the d'_{grp} listed for the six and seven-member groups are based on 400 trials or the average of two groups that completed 400 trials each.

Group performance (d'_{grp}) increased with number of group members (m), and decreased both with inter-member correlation (r) and difficulty (σ) of the conditions. These effects of m, r, and σ on group performance (d'_{grp}) are shown in Figures 6 and 7. The difficulty level of the conditions (σ) is along the x axis and d'_{grp} is plotted vertically, with data point labels representing the number of members in the group (m). Figure 6 shows the performance (d'_{grp}) for all group conditions with inter-member correlation (r) = 0.0, and Figure 7 shows the performance (d'_{grp}) for r = 0.25.

<u>Task Difficulty (σ)</u>

Similar to the effect seen in individual performance, the difficulty level (σ) had a negative effect on group

Table 7.
Average group performance (d'_{grp}) for each experimental condition, with groups of two and four members. There were 100 trials per block, with values based on 16 to 20 blocks. The unequal σ condition was based on 4 and 8 blocks.

Group Members	# in Group	Correlation (r)	Difficulty Distribution	Difficulty Condition	d'_{grp} (σ)
S_1 S_2 OR S_5 S_6	2	0.0	Equal	σ=3.0	1.02 (.33)
		0.0	Equal	σ=1.5	2.17 (.23)
		0.25	Equal	σ=3.0	0.97 (.29)
		0.25	Equal	σ=1.5	1.94 (.22)
S_1 S_2 S_3 S_4 OR S_5 S_6 S_7 S_8	4	0.0	Equal	σ=3.0	1.24 (.30)
		0.0	Equal	σ=2.0	1.97 (.37)
		0.25	Equal	σ=3.0	0.99 (.16)
		0.25	Equal	σ=2.0	1.61 (.24)
S_7 S_8 S_2 S_5	4	0.0	Unequal	σ=1.5	1.95 (.36)
		0.0	Unequal	σ=3.0	
S_2 S_5 S_7 S_8	4	0.0	Unequal	σ=1.5	1.48 (.19)
		0.0	Unequal	σ=3.0	
S_5 S_7 S_2 S_8	4	0.0	Unequal	σ=1.5	1.69 (.16)
		0.0	Unequal	σ=3.0	

Table 8.
Average group performance (d'_{grp}) for each experimental condition, with groups of six and seven members. There were 100 trials per block, with values based on 4 or 8 blocks. The values for the 8 block conditions are the average for two groups with different membership.

# in Group	Correlation (r)	Difficulty Distribution	Difficulty Condition	d'_{grp} (σ)	# of Blocks
6	0.0	Equal	σ=3.0	1.19 (.31)	8
6	0.0	Equal	σ=2.5	1.62 (.32)	4
6	0.25	Equal	σ=2.5	1.46 (.25)	4
7	0.0	Equal	σ=2.5	1.57 (.24)	8
7	0.25	Equal	σ=3.0	1.19 (.28)	8
7	0.0	Equal	σ=3.0	1.37 (.40)	4
7	0.25	Equal	σ=2.5	1.70 (.35)	4

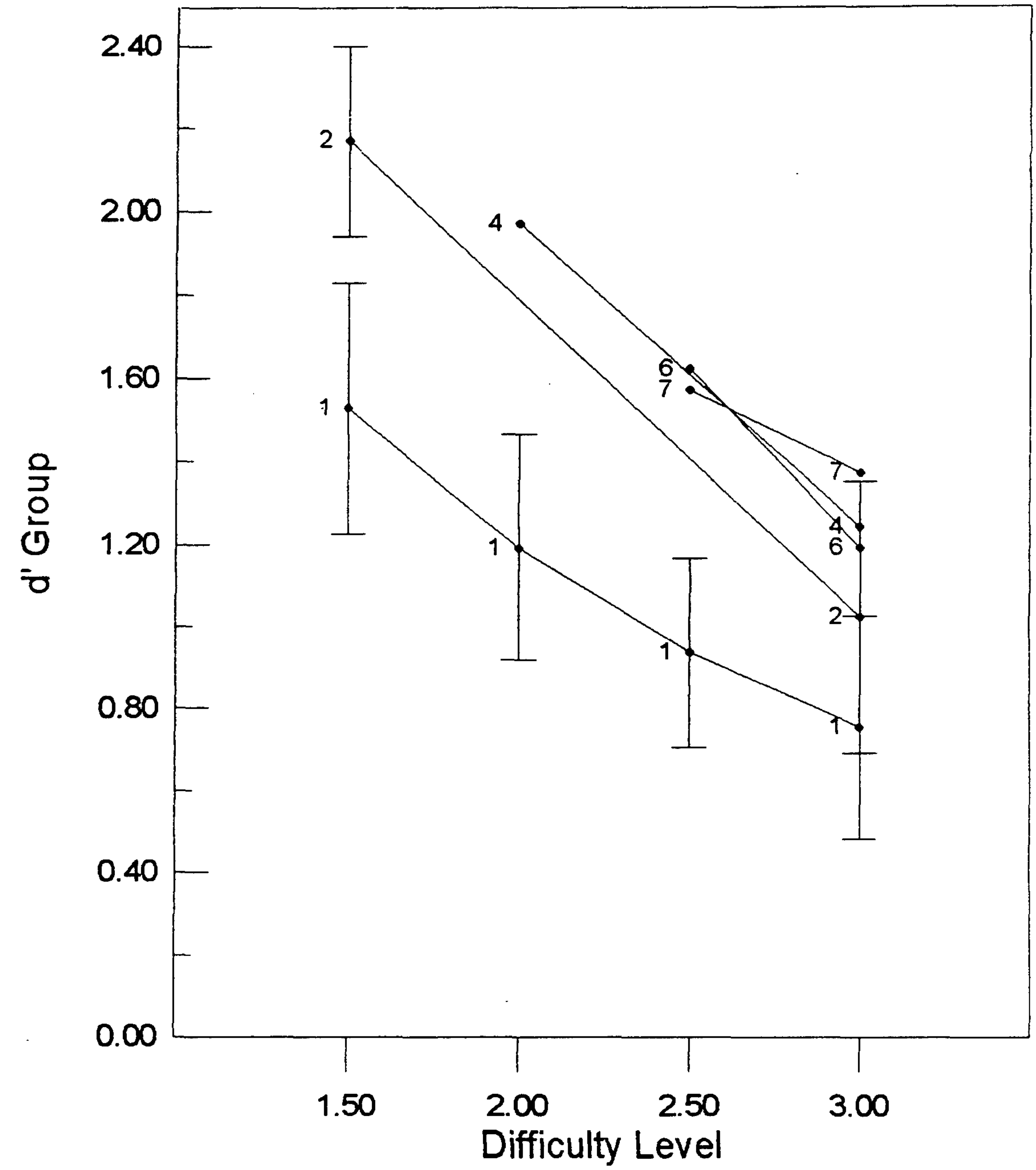


Figure 6. The group performance (d'_{grp}) plotted against difficulty level (σ) of the trials. These values are for conditions with inter-member correlation (r)=0.0. The data point labels are the number of members in the group (m). The error bars represent one standard deviation each side for the average of eight values for m=1 and all block values (19 and 16 blocks, respectively) for m=2.

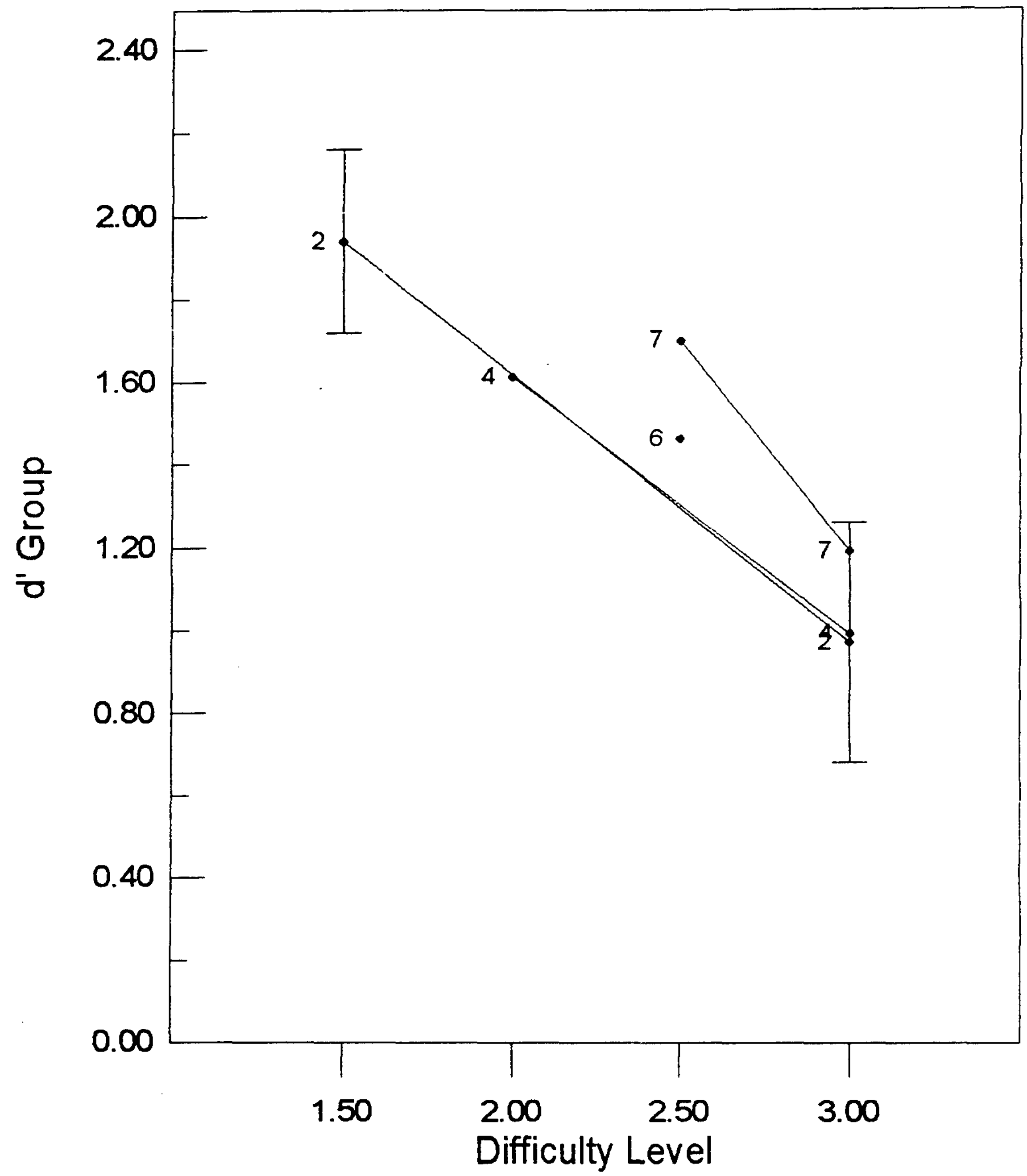


Figure 7. The group performance (d'_{grp}) plotted against difficulty level (σ) of the trials. These values are for conditions with inter-member correlation $(r)=0.25$. The data point labels are the number of members in the group (m). The error bars represent one standard deviation each side for the average of 16 and 20 blocks, respectively.

performance (d'_{grp}). This can clearly be seen in Table 7, which lists the mean performance of the two-member groups, and in Figures 6 and 7. For each size group (m), d'_{grp} decreased as the difficulty (σ) of the trials increased. The effect of difficulty (σ) on group performance (d'_{grp}) was significant in all of the two and four member, r combinations ((m=2,r=0 $F(1,33)=135.25$ $p<.001$) (m=2,r=.25 $F(1,34)=113.45$ $p<.001$) (m=4,r=0 $F(1,34)=38.80$ $p<.001$) (m=4,r=.25 $F(1,32)=75.94$ $p<.001$)). The effect of σ on d'_{grp} in six and seven member groups did not reach statistical significance ((m=6,r=0 $F(1,10)=4.26$ $p<.066$) (m=7,r=0 $F(1,10)=1.04$ $p<.332$) (m=7,r=.25 $F(1,10)=6.03$ $p<.034$)).

Number of Group Members (m)

Group performance (d'_{grp}) increased with increases in the number of group members (m). For example, Table 7 shows that in the $\sigma = 3.0$ and $r = 0.0$ condition, the average performance (d'_{grp}) for the groups with two members (S_1 S_2 & S_5 S_6) was = 1.02 and the groups of four (S_1 S_2 S_3 S_4 & S_5 S_6 S_7 S_8) had an average $d'_{grp} = 1.24$. Figures 6 and 7 graphically depict the increase in d'_{grp} as the number of group members (m) increases. This effect is fairly consistent in Figure 6 and can also be seen in Figure 8.

Figure 8 shows group performance plotted as a function of m and inter-member correlation (r), in the only difficulty condition common to several different group sizes (σ=3.0). The filled circle on the far left of the graph represents the average individual performance for all eight

subjects. The other plotted points represent the average of two groups over 8 blocks for each two or four-member group and 4 blocks for each six and seven-member group. The error bars plotted for the individual performance are based on the average for all the individual subject's performance (average of 8 values). The error bars for the group performance are based on 4 or 8 block values for six and seven-member groups and 16 to 20 values for the two and four-member groups. Only one side of the error bars were plotted to avoid overlapping lines. Within the r=0.0 condition, the positive effect of m on performance, was significant ($F(4,51)=4.56$, $p<.003$). A post-hoc (Tukey's) test showed that the individual performance was significantly less than the four and seven-member groups ($p<.05$) and was marginally less than the six-member group ($p<.06$). Within the r=0.25 condition, the increase in group performance due to increasing the number of group members from m=2 to m=7 did not reach significance ($F(2,43)=2.39$, $p<.103$).

Inter-member Correlation (r)

The effect of inter-member correlation (r) on group performance is also apparent in Figure 8 since the r = 0.0 curve is consistently higher than the r = 0.25 curve. The d'_{grp} values for the $\sigma=3.0$ condition (all groups) showed that inter-member correlation (r) had a significant ($F(1,92)=6.276$, $p<.01$) negative effect on group performance. This effect can also be seen by comparing Figure 6 (r=0.0) to Figure 7 (r=0.25).

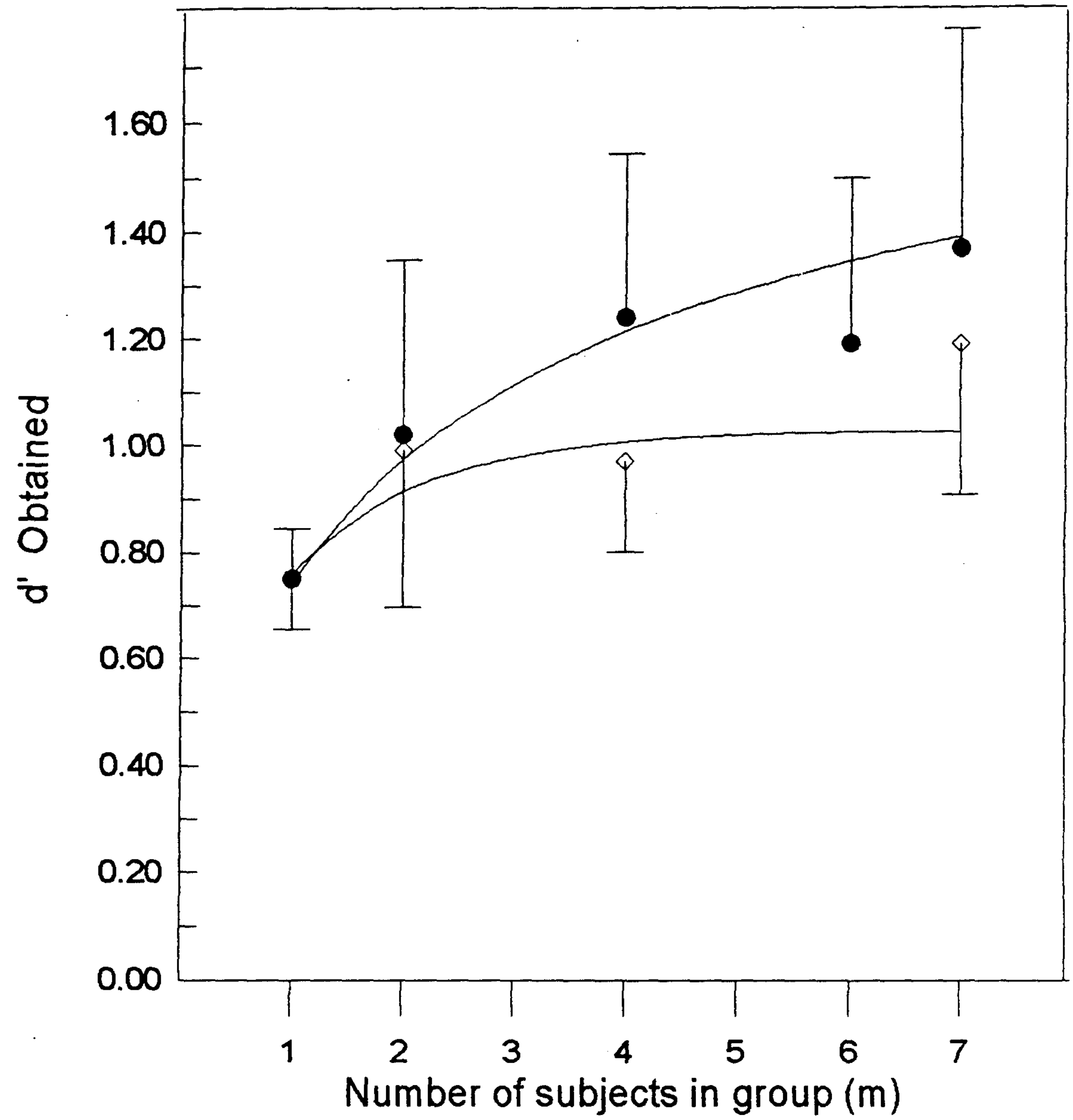


Figure 8. Best fit lines for the (σ)=3.0 condition. The correction from the ideal predictions to actual performance is inversely proportional to the number of group members and is described in the text. The number of group members (m) is along the x axis and d' along the y axis. The solid circles represent the independent (r=0.0) condition, and the open diamonds represent the correlated (r=0.25) condition. The error bars represent one standard deviation with the values based on 4 or 8 blocks in m=6, m=7 or 16-20 blocks for the m=2, m=4 groups.

Group performance fit to ideal model. The smooth functions in Figure 8 were calculated using a least squares fit (Scientist Plot) of the data to equation 6 (Sorkin & Dai, 1994), and assuming an additional source of noise, σ^2_G. Performance was fit to the following equation:

$$d'_{grp} = \left[\frac{1}{1 + m\,\sigma^2_G}\right]^{\frac{1}{2}} \left[\frac{m \operatorname{Var}(d')}{1 - r} + \frac{\bar{m}\,(d')^2}{1 + r\,(m-1)}\right]^{\frac{1}{2}} . \qquad (28)$$

Where m is the number of group members and σ^2_G represents the additional noise source. The second part of the equation is the original ideal group equation, given in equation 6 (Sorkin & Dai, 1994). The effect of the additional noise is to reduce performance somewhat as the group size increases. In this fit, $\sigma^2_g = 0.20$.

Group Weighting Strategy

In order to determine the relative weights assigned to each member during the group process, Lutfi's (in press) correlational weighting analysis was utilized. The procedure used to derive the relative weights is based on the point-biserial correlation between the stimulus presented to each group member and the group response, scaled by the variability of the presented information. The weighting procedure is explained in the Appendix. Equation 9 (from Sorkin & Dai, 1994), described earlier, was used to determine the ideal weights for each member depending on their individual performance and the trial condition. The actual relative weights and the ideal weights are shown in Figures

9 through 16, with the relative weights represented by the open bars and the ideal weights represented by the hashed bars.

The unequal σ conditions show that the groups can differentiate differences in individual performance and assign weights that are similar to ideal. For example, in the unequal σ group trials shown in Figure 13, the group members that received information that was difficult to discriminate (σ=3.0), clearly received less weight than the other two group members. All three subject-difficulty combinations tested show this effect.

The weights assigned during the equal σ condition trials appear to be somewhat variable. The weights do not appear to follow a specific pattern, but the differences in individual performance are very small to begin with.

However, there were clearly some deviations from a weighting strategy based on individual performance. An example of this is depicted in the negative weighting of S_4 on Figure 16. S_4 forced her group to wait for her because she was late for two consecutive experimental sessions. This seems to have had a negative effect on the relative weight that the group assigned to her.

<u>Differences Between Actual and Ideal Weights</u>

In order to test the effectiveness of the group weighting strategy, we tested the significance of the difference between actual and ideal weights. First, the correlations between each group members' stimuli (sum of stimulus gauges)

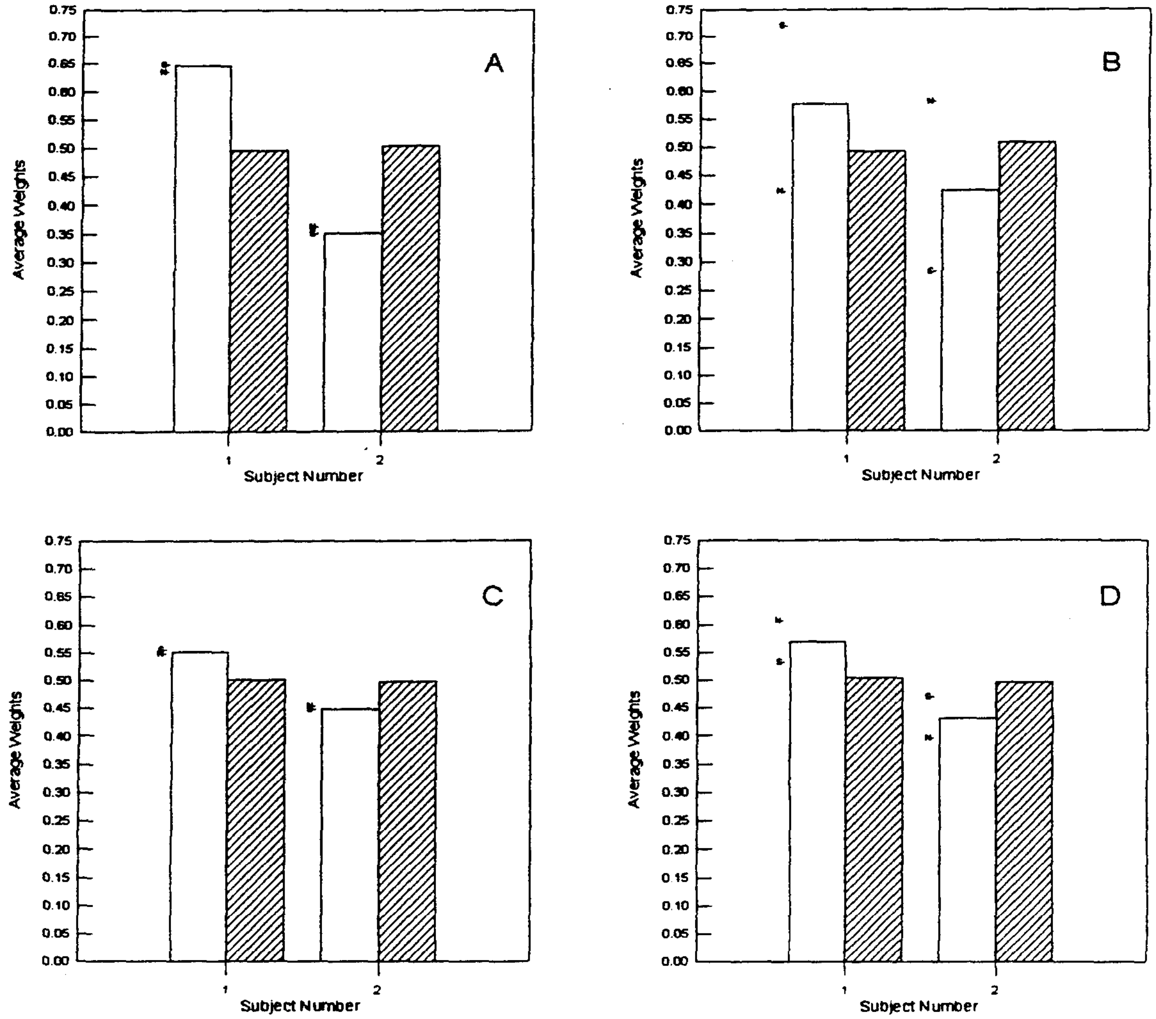


Figure 9. The relative weights for a two-member group. The average relative weight is represented by the open bar and the ideal weights are represented by the hashed bar. A represents the $(\sigma)=1.5$, $r=0.0$ condition, B represents the $(\sigma)=1.5$, $r=.25$ condition, C represents the $(\sigma)=3.0$, $r=0.0$ condition, and D represents the $(\sigma)=3.0$, $r=.25$ condition. The S- symbol represents the signal trials and the N- symbol represents the noise trials (400-600 each).

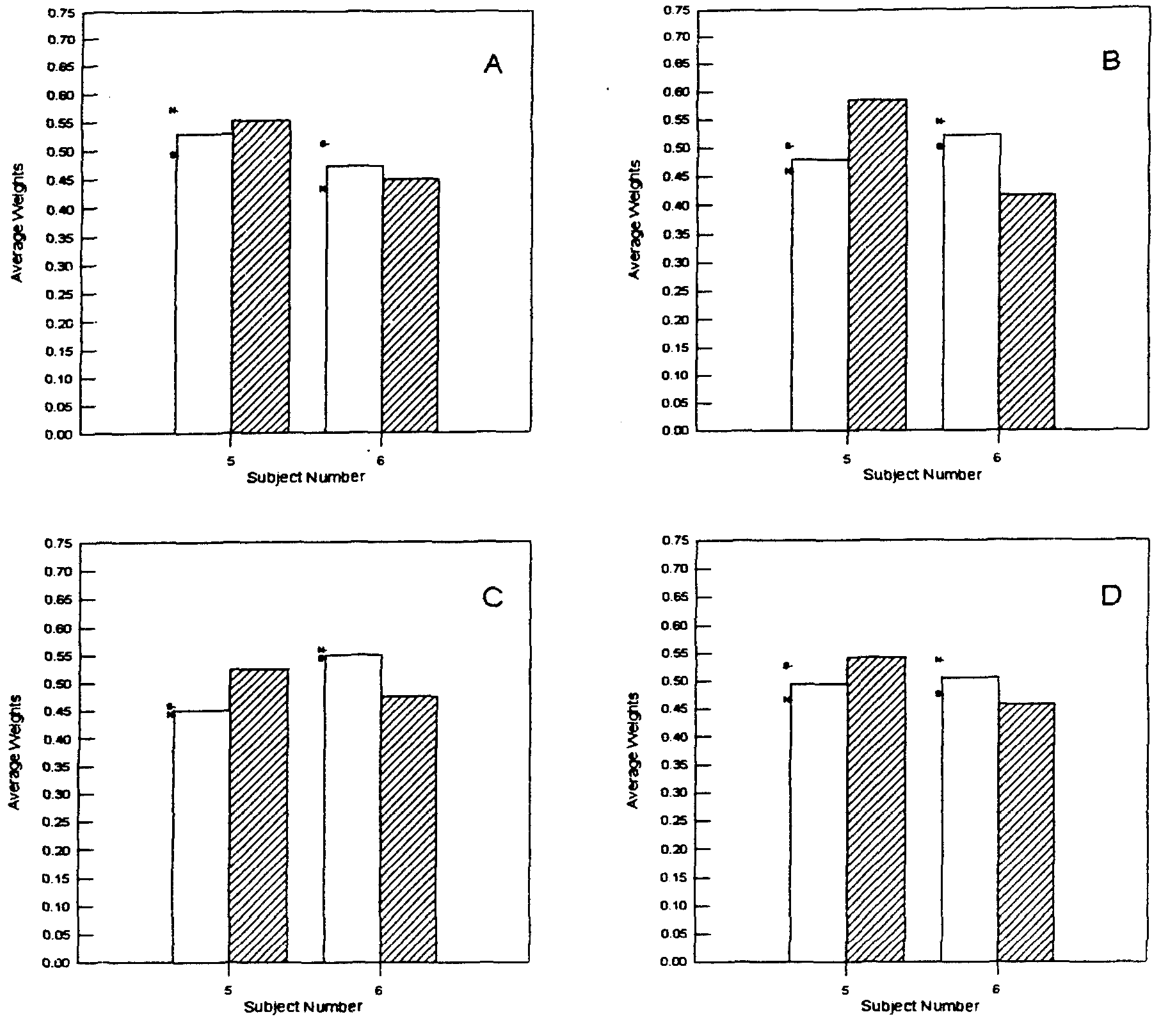


Figure 10. The relative weights for a two-member group. The average relative weight is represented by the open bar and the ideal weights are represented by the hashed bar. A represents the (σ)=1.5, r=0.0 condition, B represents the (σ)=1.5, r=.25 condition, C represents the (σ)=3.0, r=0.0 condition, and D represents the (σ)=3.0, r=.25 condition. The S- symbol represents the signal trials and the N- symbol represents the noise trials (400-600 each).

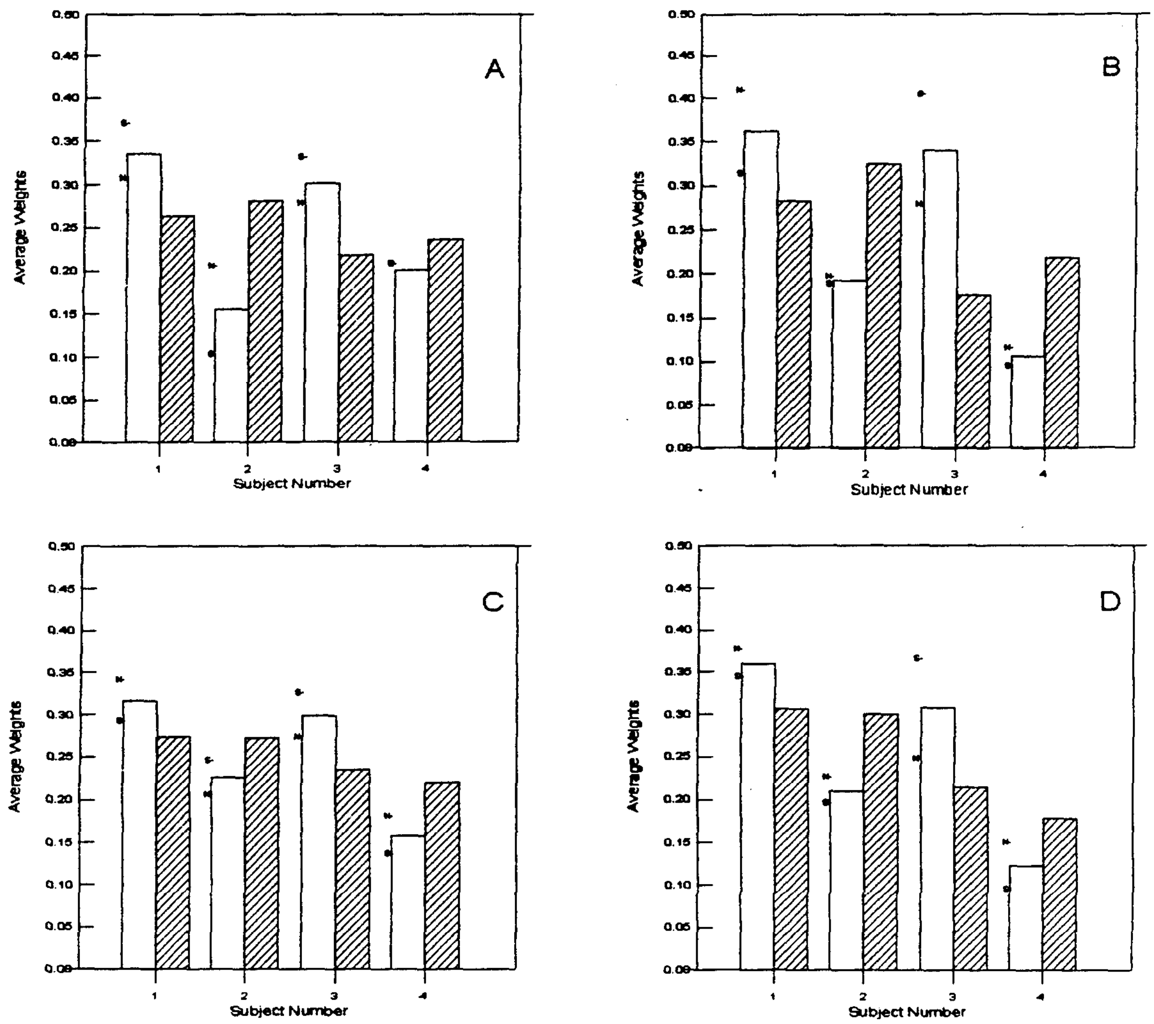


Figure 11. The relative weights for a four-member group. The average relative weight is represented by the open bar and the ideal weights are represented by the hashed bar. A represents the (σ)=2.0, r=0.0 condition, B represents the (σ)=2.0, r=0.25 condition, C represents the (σ)=3.0, r=0.0 condition, and D represents the (σ)=3.0, r=0.25 condition. The S- symbol represents the signal trials and the N- symbol represents the noise trials (400 each).

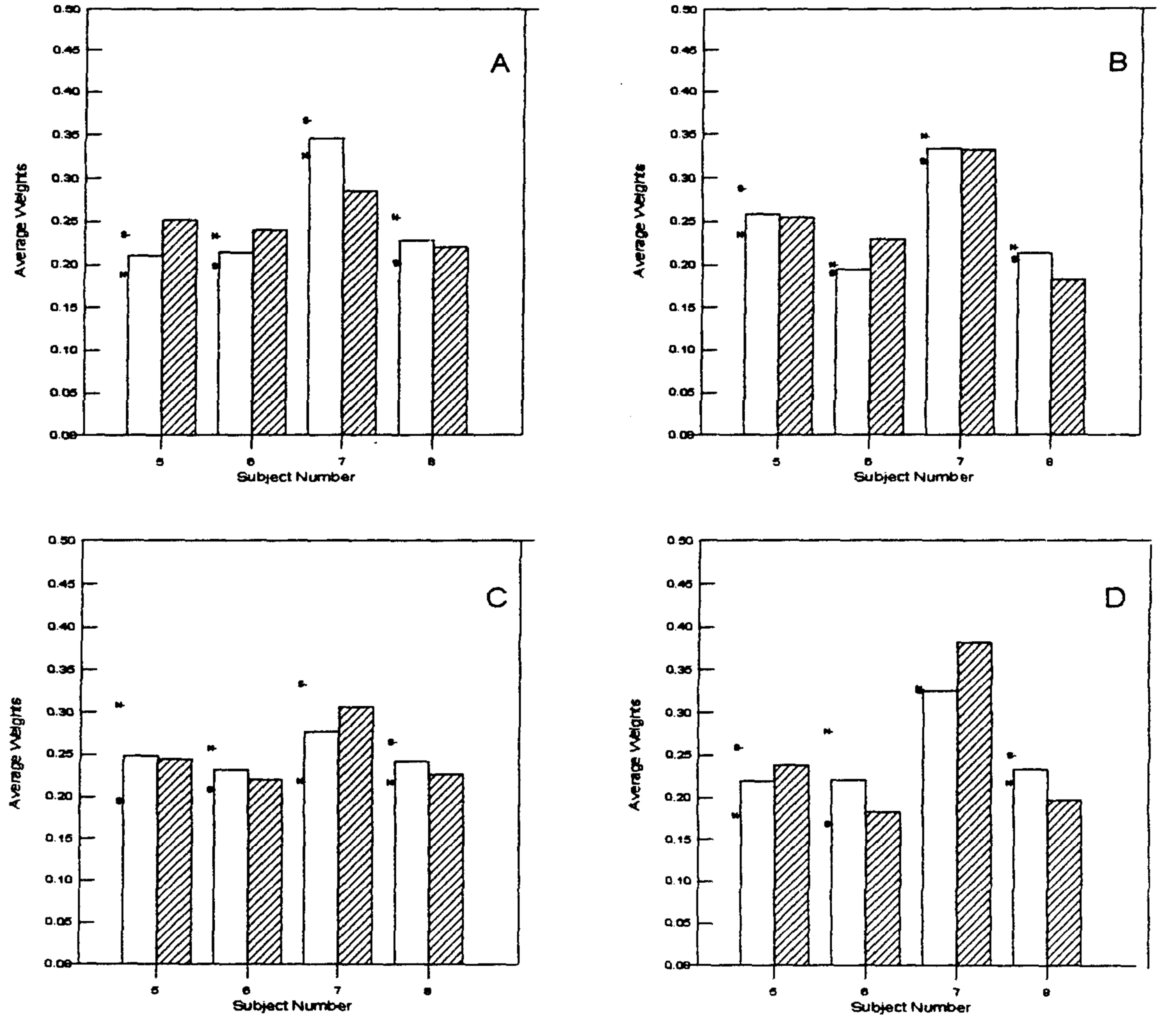


Figure 12. The relative weights for a four-member group. The average relative weight is represented by the open bar and the ideal weights are represented by the hashed bar. A represents the (σ)=2.0, r=0.0 condition, B represents the (σ)=2.0, r=0.25 condition, C represents the (σ)=3.0, r=0.0 condition, and D represents the (σ)=3.0, r=0.25 condition. The S- symbol represents the signal trials and the N- symbol represents the noise trials (400 each).

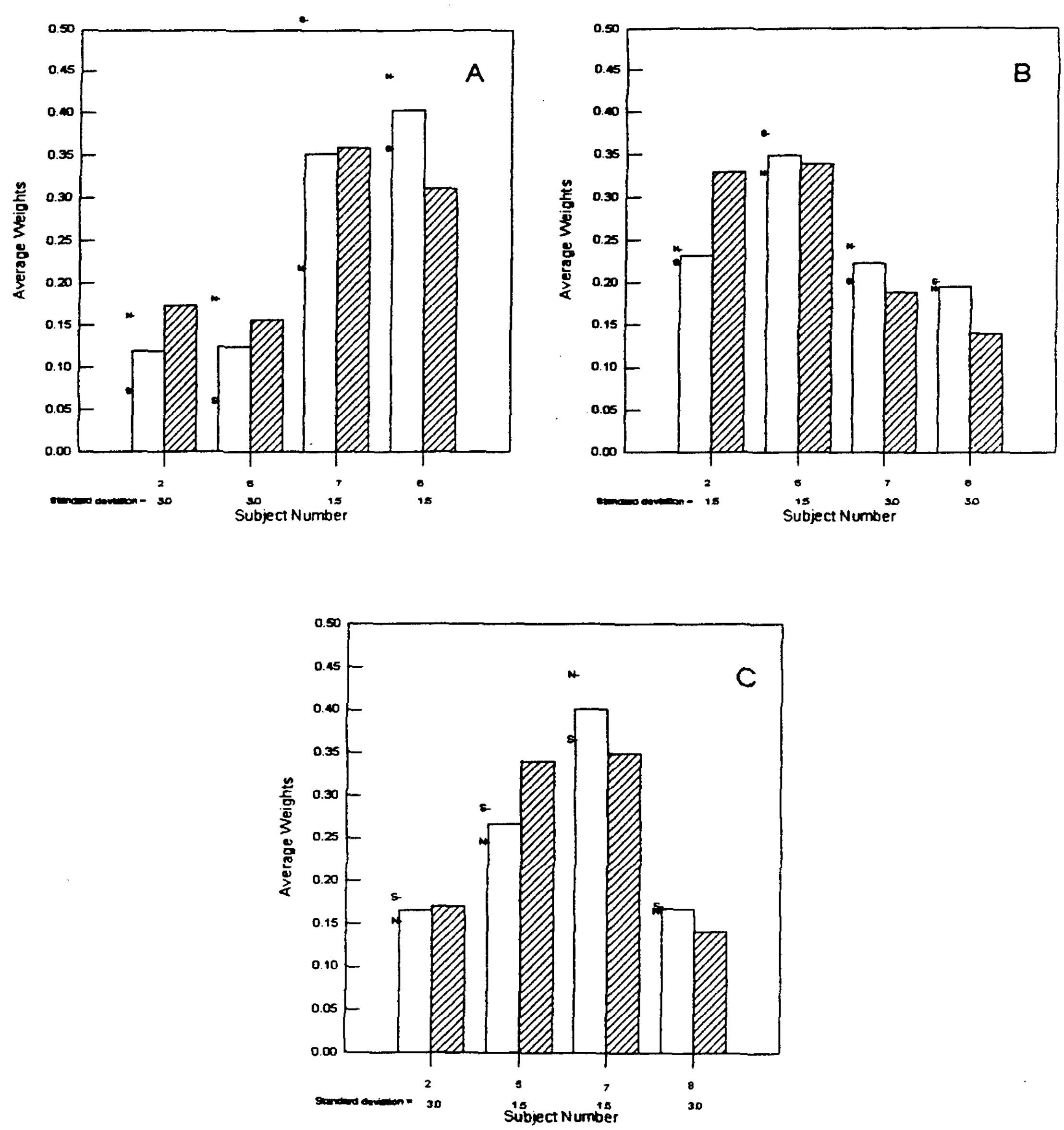


Figure 13. The relative weights for a four-member group under the unequal difficulty (σ) condition. A and C are based on 400 trials and B is based on 800 trials. The S- symbol represents the signal trials and N- represents the noise trials. The standard deviation (difficulty) of the presented stimuli distributions (single gauge) are listed beneath the subject number.

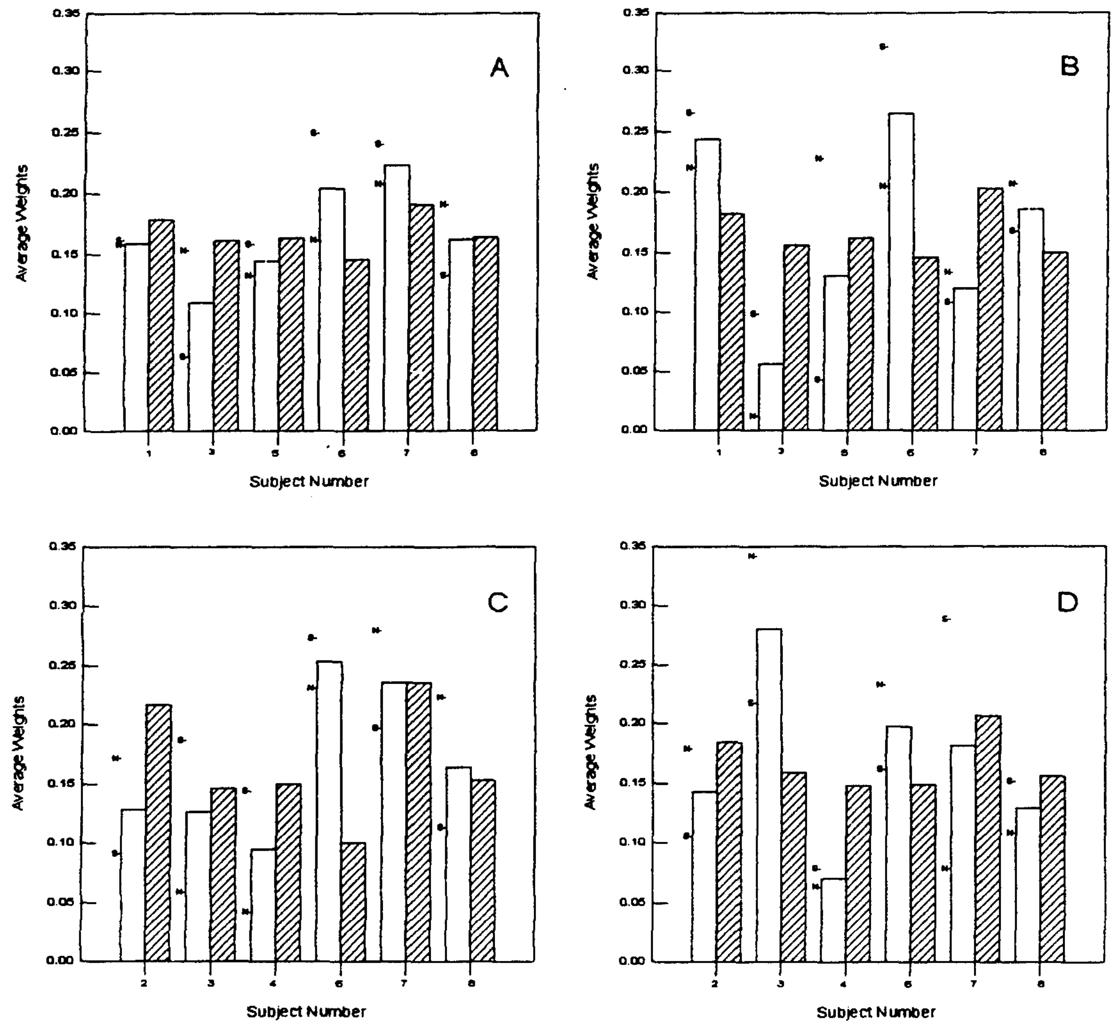


Figure 14. The relative weights for a six-member group. The average relative weights are represented by the open bar and the ideal weights are represented by the hashed bar. A represents the (σ)=2.5, r=0.0 condition, B represents the (σ)=3.0, r=0.0 condition, C represents the (σ)=2.5, r=0.25 condition, and D represents the (σ)=3.0, r=0.25 condition. The S- symbol represents the signal trials and the N- symbol represents the noise trials. Each graph is based on 400 trials.

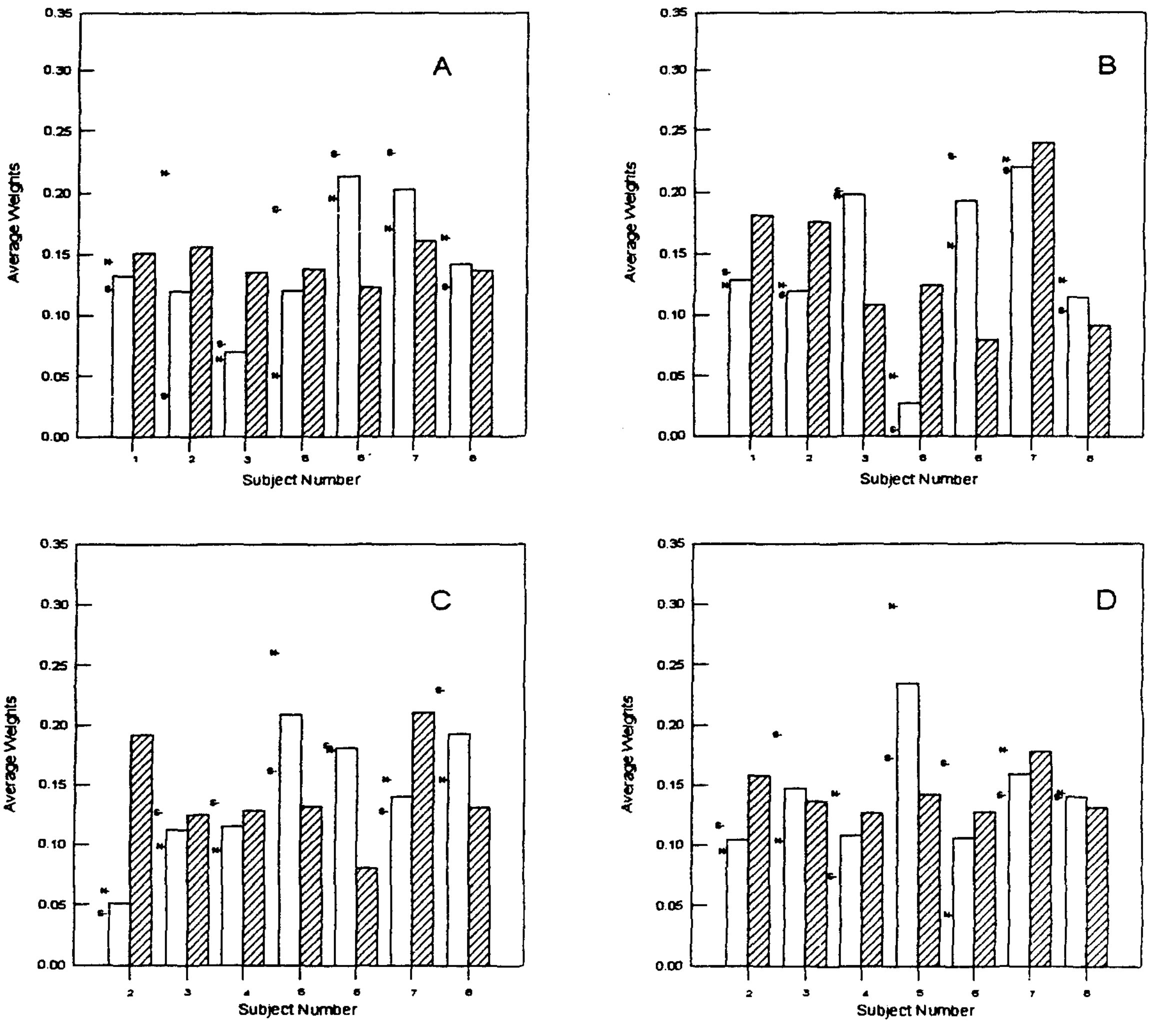


Figure 15. The relative weights for a seven-member group. The average relative weights are represented by the open bar and the ideal weights are represented by the hashed bar. A represents the $(\sigma)=2.5$, $r=0.0$ condition, B represents the $(\sigma)=3.0$, $r=0.25$ condition, C represents the $(\sigma)=2.5$, $r=0.25$ condition, and D represents the $(\sigma)=3.0$, $r=0.0$ condition. The S- symbol represents the signal trials and the N- symbol represents the noise trials. Each graph is based on 400 trials.

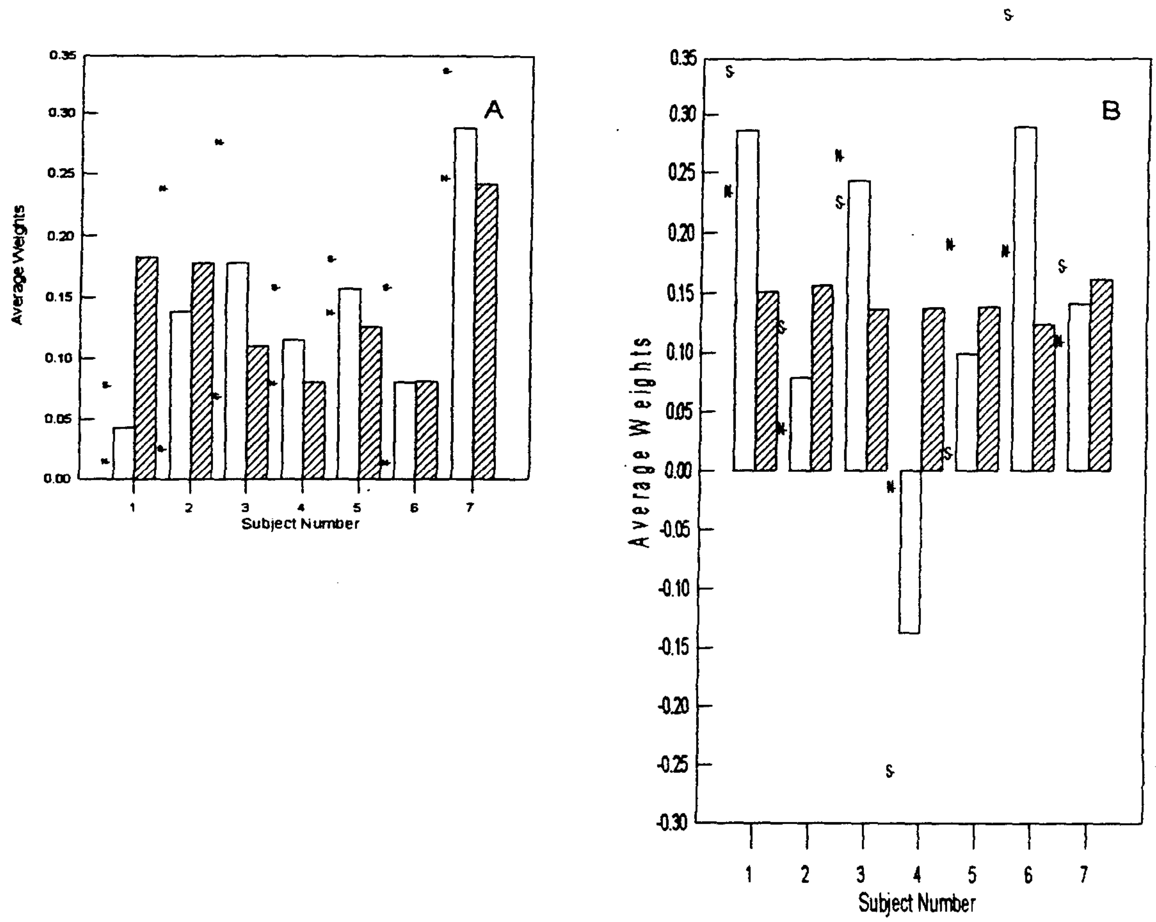


Figure 16. The relative weights for a seven-member group. The average relative weights are represented by the open bar and the ideal weights are represented by the hashed bar. A represents the (σ)=3.0, r=0.25 condition, and B represents the (σ)=2.5, r=0.0 condition. The S- symbol represents the signal trials and the N- symbol represents the noise trials. Each graph is based on 400 trials .

and the group response were calculated. These correlations were then converted to relative weights (Appendix). Ideal weights ($\hat{a}_i$) were derived by using equation 9.

In order to determine if the relative weights (a_i) were statistically different from the ideal weights, a 99% confidence interval around the relative weight was formed.

$$a_i \pm 2.575(\sigma_{ai}) \tag{29}$$

The standard deviation of the actual weights (σ_{ai}) was determined as follows: First, the correlations between the sum of the stimulus displays and the group response (r_{Rxi}) were transformed into Fisher's Z scores (Z_r).

$$Z_r = (1/2) \log_e \left[\frac{1 + r_{Rxi}}{1 - r_{Rxi}}\right] \tag{30}$$

The Fisher's Z has an estimated standard error (σ_{Zr})= $1/\sqrt{n-3}$ for standard correlations. [When the assumption of independent information sources is violated (i.e. r =.25), partial correlations can be utilized. The estimated standard error of Z is equal to $1/\sqrt{n-4}$, or $1/\sqrt{n-(m+3)}$ for higher order correlations, where n is the number of trials and m is the number of variables being partialled out (Thorndike, 1978)]. Next, the ratio $(\sigma_{Zr})/Z_{rRxi}$ was set equal to $(\sigma_{rRxi})/r_{Rxi}$ in order to determine an estimated standard deviation of the correlation (σ_{rRxi}).

$$\frac{\sigma_{Zr}}{Z_{rRxi}} = \frac{\sigma_{rRxi}}{r_{Rxi}} \tag{31}$$

Given that the relative weights (a_i), (in the equal difficulty condition) can be thought of as the stimulus-response

correlation (r_{Rxi}) times a scaling constant, the equality given in equation 32 was used to determine σ_{ai}.

$$\frac{\sigma_{ai}}{a_i} = \frac{\sigma_{rRxi}}{r_{Rxi}} \tag{32}$$

These tests showed σ_{ai} = .035 - .065.

Examination of all the group conditions, within the equal difficulty condition, showed 5 out of a possible 29 conditions where the groups assigned a weight that was significantly different than ideal ($p<.01$). It was interesting that S_2 was weighted significantly less than ideal on two out of four conditions in the four member group ($\sigma=2.0$, $r=0$ and $r=.25$, Figure 11 A & B) and once in the two-member group ($\sigma=1.5$, $r=0.0$, Figure 9 A). The weights in the unequal difficulty condition (Figure 13) showed very small variations from ideal weights; none exceeded the criterion. These findings show that the unequal σ condition displayed nearly ideal weights and the equal σ condition displayed weights that were more variable but not drastically different from ideal.

To be sure that the weighting test described above was appropriate, a Monte Carlo simulation was conducted. The simulation used a group of four members in a trial where the ideal weights were equal. The simulated group decision statistic (Z from Figure 1) was based on equal weighting of all members and the criterion (Z_c) was placed midway between signal and noise means. The sum of the stimulus displays for each group member and the group response were recorded.

The simulated trials were then separated and analyzed in signal and noise trials. The trials were analyzed according to Lutfi's correlational analysis (Appendix) in order to determine the simulated ideal weights. The mean and standard deviation of a single "ideal member's" normalized weight based on 500 blocks with 100 trials per block was .257 and .08 respectively. Therefore, the estimated standard error of a_i is

$$\sigma_{ai} = \frac{.08}{\sqrt{n}} = .04 \qquad (33)$$

where n is the number of blocks of trials. This value (.04 for four trial blocks) was very similar to the σ_{ai} values (.035 - .065) calculated in equation 32. Consequently, it can be concluded that the method of testing the relative weights, described earlier, was appropriate.

Weights When Group Responder versus Non-responder

The effect of being the responder for the group was examined by evaluating a representative four-member (S_1 - S_4) group's performance under the σ=2.0, r=0.0 condition. Within noise or signal trials, the correlations between each subject's stimulus (sum of the presented stimulus gauge values) and the group response were calculated for two cases. The correlations were calculated for trials when she/he responded for the group and for those when she/he was a non-responding member of the group. In the equal σ condition, these correlations are proportional to the relative weights.

The relative weight for a subject was consistently higher when she/he responded for the group compared to when the subject did not respond. This can be seen in Figure 17. The weights shown in Figure 17 represent each subjects' relative weight when assigned to be a responder (open bar) and non-responder (hashed bar) under signal and noise trials. The error bars represent estimates of the standard error of the mean, $\sqrt{1/n-3}$. Where n represents the number of trials.

The correlations were converted into Z-values using Fisher's Z transformation (equation 30). The differences between these values were then used to determine if the members' relative weights were different when they responded for the group than when they did not respond for the group. Since the number of trials was different for each case (approximately 100 responder, 300 non-responder), the form of the test was:

$$Z = \frac{Z_{responder} - Z_{non\text{-}responder}}{\sqrt{1/(n_{responder} - 3) + 1/(n_{non\text{-}responder} - 3)}} \quad (34)$$

Where the Z values are for a single subject (Thorndike, 1978). Since it is possible that the group assigned less or more weight to the responder, a two-tail 95% confidence interval was used to determine if the correlations were significantly different. Although all of the eight tests showed that more weight was assigned to the subject when acting as a responder versus a non-responder, none of these were statistically significant ($p>.05$).

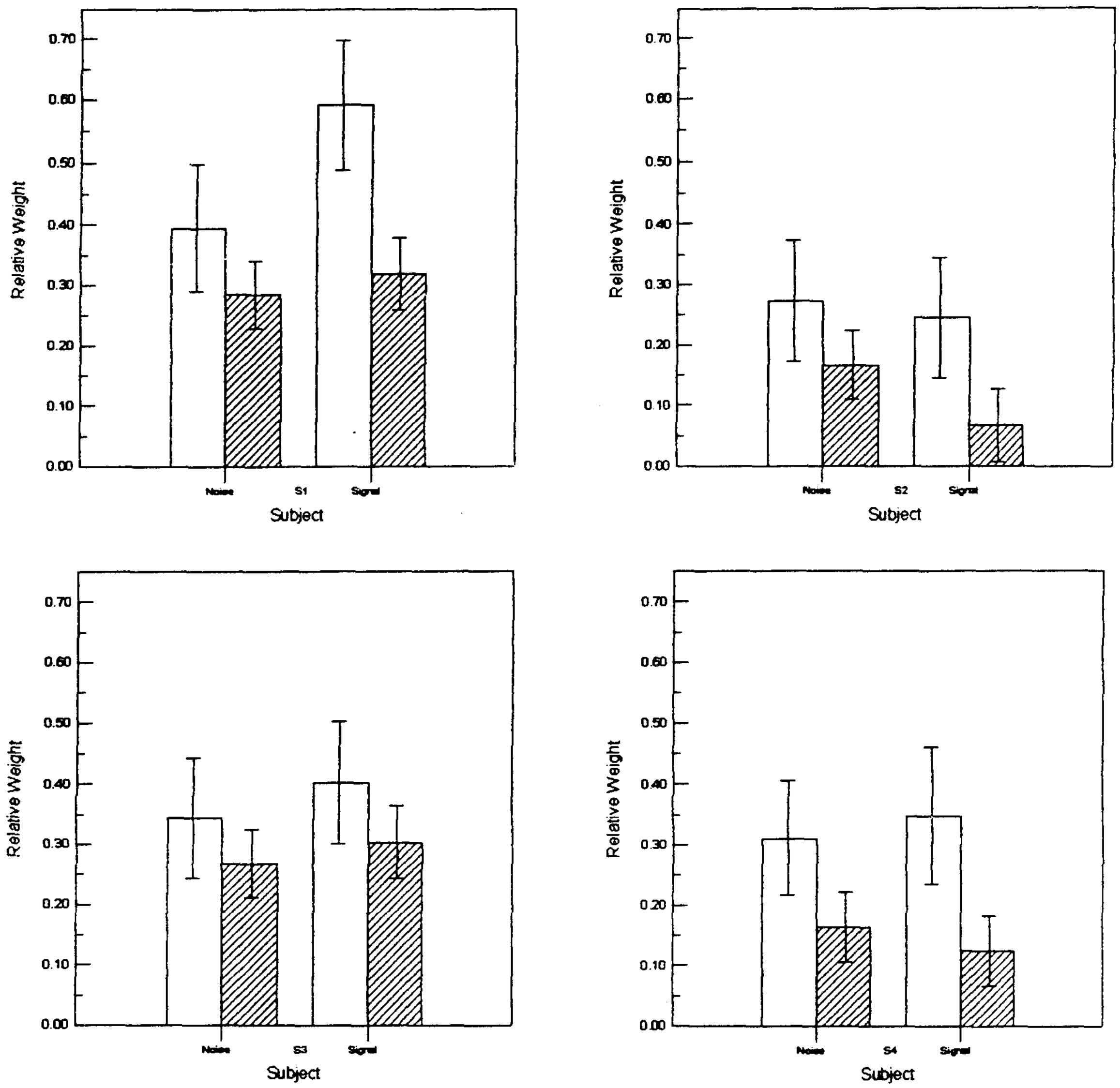


Figure 17. Relative weights assigned to members of a four-member group when chosen to respond for the group (open bar) and when a non-responder (hashed bar). The weights are based on approximately 400 trials for both signal and noise trials divided into approximately 100 trails as responder and 300 trials when non-responder. The group condition (σ=2.0, r=0.0) was randomly selected as a representative example. The error bars represent the estimated standard error of the mean for $\sqrt{1/n-3}$.

Notes on member's interaction tendencies. One possibility for some of the deviations from ideal weights is that the groups used one or more social attributes as weighting criteria. During the group trials, several of the group interactions were monitored by the experimenter. Table 9 depicts the tendencies for individual subjects to interact during the group trials, based on notes taken during these interactions. This is a somewhat subjective measurement but it may account for the consistent lower relative weights assigned to S_2 even though she was one of the best individual performers.

On the opposite end of Table 9, S_8 was highly interactive. S_8 received weights slightly larger than ideal on all trial conditions in all of the four-member group interactions in which she participated. These examples suggest that social factors may have affected the relative weights assigned to individual subjects during the group decision process.

Efficiency of Group Performance Across Conditions

Groups performed relatively effectively across trial conditions. However, their performance, relative to the ideal, tended to decrease as m increased. Figure 18 shows the obtained group performance (d'_{grp}) plotted against ideal group performance (d'_{ideal}) for all r and σ conditions. The numbers plotted on the figure indicate the size of the groups.

Table 9.
Individual group member's tendencies toward interacting during group interaction. These tendencies were estimated from experimenter's notes taken during group interaction.

Interaction Tendencies

Non-Verbal Quiet/Reserved	Moderately Interactive	Highly Verbal Loud/Interactive
S_2, S_6	S_1, S_3, S_4	S_5, S_7, S_8

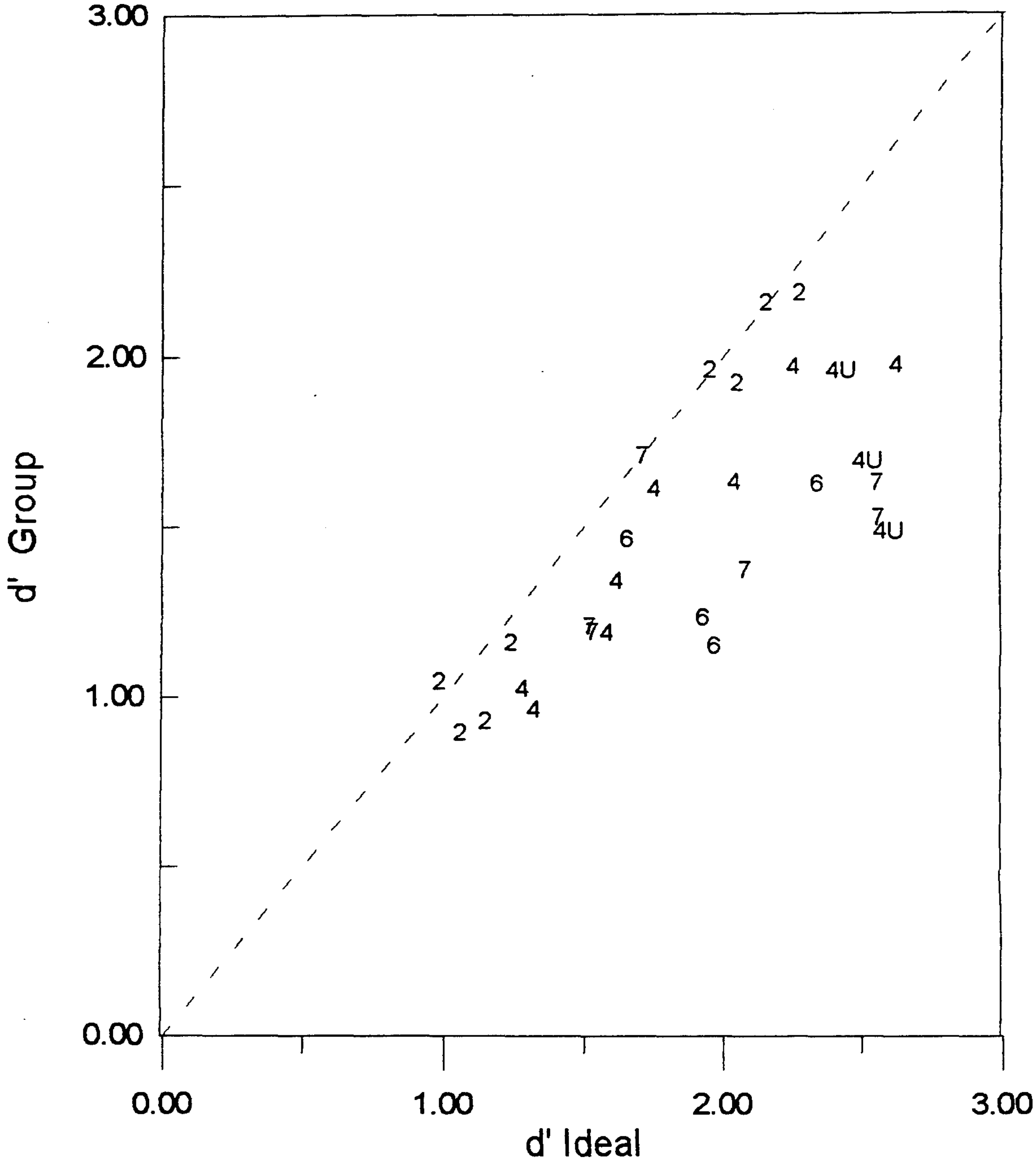


Figure 18. Group performance (d'_{grp}) plotted against (d'_{Ideal}). All difficulty and inter-member correlation conditions are displayed. The data point labels are the size of the group (m). The dashed line along the diagonal corresponds to an actual group performing at d'Ideal.

As m increases, the distance from ideal performance (major diagonal) appears to increase. The groups appear to be clustered equidistant from the diagonal based on the group size. For example, the two-member groups are all very close to the diagonal. The four-member groups are farther away from the diagonal but are all approximately the same distance from the diagonal. The six-member and seven-member groups are consistently the farthest from the diagonal. The only exceptions were two of the unequal σ (4U) group conditions; these were similar to the larger groups in their relative performance.

Group Efficiencies (^{n}weight and ^{n}obs)

The group observed efficiencies ($^{n}{}_{obs}$) and the relative weighting efficiencies ($^{n}{}_{weight}$) for all group conditions are listed in Tables 10 and 11. Overall efficiency ($^{n}{}_{obs}$) was obtained by squaring the ratio $d'_{obtained}/d'_{Ideal\ Group}$ (Tanner & Birdsall, 1958), while weighting efficiency ($^{n}{}_{weight}$) was obtained by squaring the ratio $d'_{weight}/d'_{Ideal\ Group}$ (Berg, 1991). The value d'_{weight} assumes ideal group performance except it utilizes the actual weights (a_i) assigned by the groups.

Overall efficiency ($^{n}{}_{obs}$) in the equal σ conditions decreased ($F(3,22)=6.615$ $p<.002$) with group size (m). A post-hoc analysis (Tukey's test), showed that the efficiency of the two-member groups was significantly higher than all the other groups ($p<.05$). Group weighting efficiency was high across all conditions, but marginally decreased from

Table 10.
Efficiency measures for group performance in each experimental condition, with groups of two and four members. There were 100 trials per block, with values based on 8 or 12 blocks.

Group Members	# in Grp	r	Difficulty Distribution	Difficulty Condition	Efficiency Measures n_{obs}	n_{weight}	n_{noise}
S_1 S_2	2	0.0	Equal	σ=3.0	0.871	0.886	0.984
S_1 S_2	2	0.0	Equal	σ=1.5	0.923	0.895	1.030
S_1 S_2	2	0.25	Equal	σ=3.0	0.653	0.825	0.792
S_1 S_2	2	0.25	Equal	σ=1.5	0.879	0.946	0.929
S_1 S_2	4	0.0	Equal	σ=3.0	0.684	0.848	0.807
S_3 S_4	4	0.0	Equal	σ=2.0	0.566	0.866	0.654
S_1 S_2	4	0.25	Equal	σ=3.0	0.531	0.744	0.714
S_1 S_2	4	0.25	Equal	σ=2.0	0.638	0.838	0.762
S_5 S_6	2	0.0	Equal	σ=3.0	0.707	0.862	0.820
S_5 S_6	2	0.0	Equal	σ=1.5	1.004	0.974	1.030
S_5 S_6	2	0.25	Equal	σ=3.0	1.124	0.814	1.378
S_5 S_6	2	0.25	Equal	σ=1.5	1.009	0.931	1.083
S_5 S_6	4	0.0	Equal	σ=3.0	0.563	0.927	0.608
S_7 S_8	4	0.0	Equal	σ=2.0	0.766	0.974	0.787
S_5 S_6	4	0.25	Equal	σ=3.0	0.639	0.823	0.776
S_5 S_6	4	0.25	Equal	σ=2.0	0.842	0.930	0.906
S_7 S_8	4	0.0	Unequal	σ=1.5	0.652	1.026	0.635
S_2 S_5	4	0.0	Unequal	σ=3.0			
S_2 S_5	4	0.0	Unequal	σ=1.5	0.325	0.908	0.358
S_7 S_8	4	0.0	Unequal	σ=3.0			
S_5 S_7	4	0.0	Unequal	σ=1.5	0.450	1.030	0.437
S_2 S_8	4	0.0	Unequal	σ=3.0			

Table 11.
Efficiency measures for group performance in each experimental condition, with groups of six and seven members. There were 100 trials per block, with values based on 4 blocks.

Group Members	# in Grp	r	Difficulty Distribution	Difficulty Condition	Efficiency Measures n_{obs}	n_{weight}	n_{noise}
S_1 S_3 S_5 S_6 S_7 S_8	6	0.0	Equal	σ=3.0	0.339	0.742	0.457
		0.0	Equal	σ=2.5	0.482	0.919	0.524
S_2 S_3 S_4 S_5 S_7 S_8	6	0.0	Equal	σ=3.0	0.406	0.792	0.513
		0.25	Equal	σ=2.5	0.779	0.817	0.953
S_1 S_2 S_3 S_5 S_6 S_7 S_8	7	0.0	Equal	σ=2.5	0.353	0.862	0.410
		0.25	Equal	σ=3.0	0.626	0.693	0.903
S_2 S_3 S_4 S_5 S_6 S_7 S_8	7	0.0	Equal	σ=3.0	0.434	0.842	0.516
		0.25	Equal	σ=2.5	0.994	0.779	1.275
S_1 S_2 S_3 S_4 S_5 S_6 S_7	7	0.0	Equal	σ=2.5	0.403	0.478	0.843
		0.25	Equal	σ=3.0	0.603	0.692	0.871

m=1 to m=7 ($F(4,26)=4.59$ $p<.012$). Figure 19 depicts the average observed and weighting efficiencies across all equal σ conditions (all σ levels and r values) plotted against the number of group members (m). Figure 19 shows a decrease in observed efficiency with an increasing number of group members (m). The two-member groups performed at levels around 0.90 while the efficiency of the larger groups decreased to levels around 0.60. The error bars represent the standard deviations for all of the trial block efficiency measures. Table 12 lists the average observed and weighting efficiencies for each group size and difficulty distribution.

There were no large effects on overall or weighting efficiency due to difficulty(σ), or inter-member correlation (r). For n_{obs} ((σ, $F(4,26)=3.11$ $p<.047$) (r, $F(2,26)=4.27$ $p<.05$)) and for n_{weight} ((σ, $F(4,26)=3.39$ $p<.036$) (r, $F(2,26) =.445$ $p<.511$)).

Table 12 lists the averages of all of the efficiency measure values as well as the values for the unequal σ condition. The unequal σ condition displayed very close to ideal weighting efficiency (0.988). The weighting efficiency measures varied very little across conditions (standard deviation = 0.122). This was equal to about half of the variability of the overall efficiency (standard deviation = 0.223). The values of the standard deviations of these efficiencies across the different group sizes are shown in Table 12.

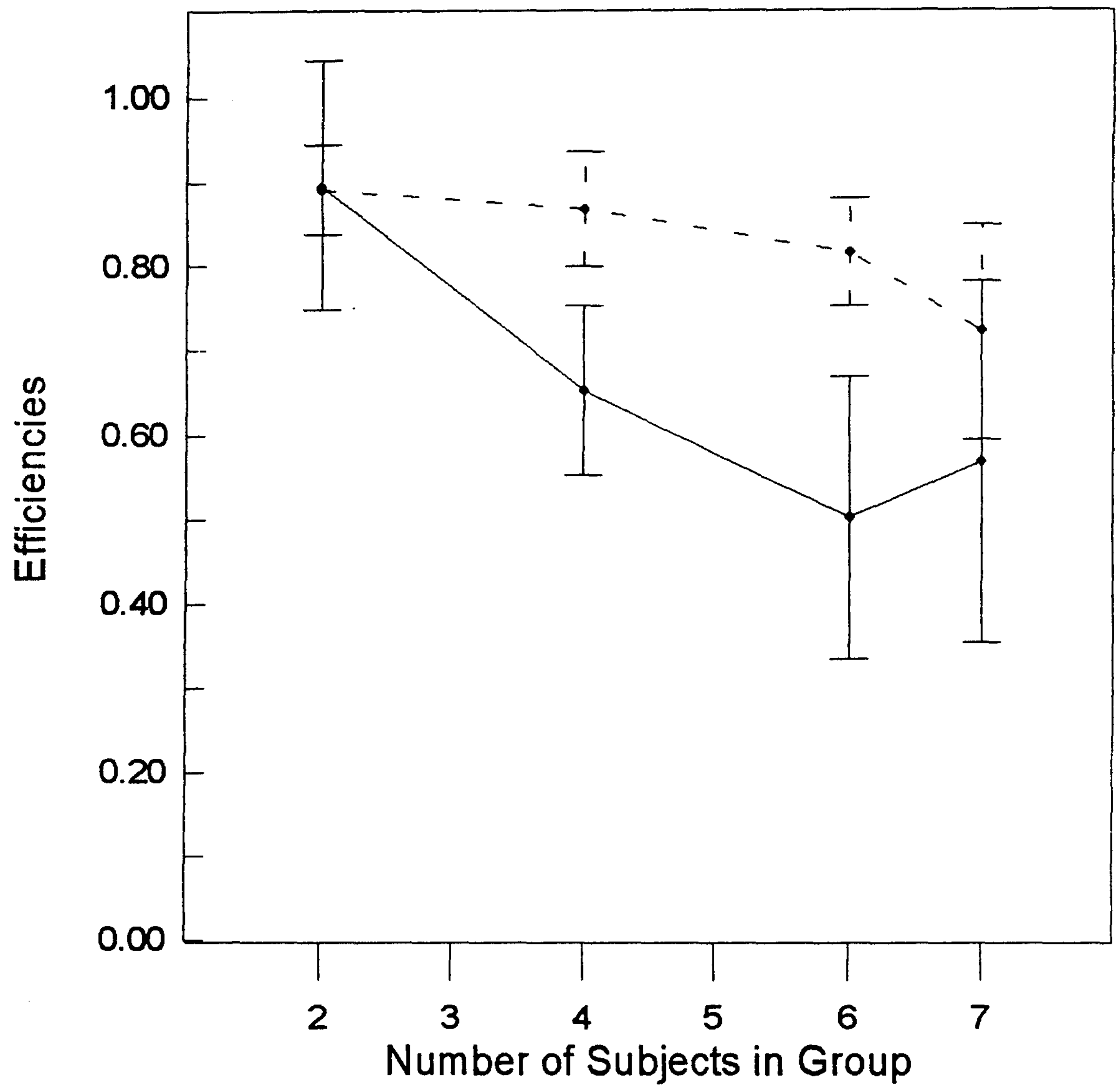


Figure 19. Efficiency measures plotted versus group size (m). Observed efficiency (n_{obs}) for all equal difficulty conditions is plotted with the solid line. Weighting efficiency (n_{weight}) is shown by the dotted line. Standard deviations of the efficiency measures are also plotted. The plotted values are based on 12 to 32 blocks of trials.

Table 12.
Efficiency measures for group performance averaged over each experimental condition. All of the experimental conditions for a given size group were used to determine the average and standard deviations for each these measures.

# in Grp	Difficulty Distribution	Mean Efficiency Measures n_{obs} (σ)	n_{weight} (σ)
2	Equal	0.896 (.147)	0.892 (.053)
4	Equal	0.654 (.100)	0.869 (.068)
4	Unequal	0.476 (.135)	0.988 (.057)
6	Equal	0.502 (.168)	0.818 (.065)
7	Equal	0.569 (.215)	0.724 (.128)

DISCUSSION

Individual group members performed consistently across trial conditions and throughout the study. This shows that measured group performance was due to changes in the experimental conditions and not to fluctuations in individual performance. The difficulty (σ) of the trials directly affected both individual and group performance. As the standard deviation of the presented stimuli (σ) distributions became larger, performance (d') declined in the predicted manner, consistent with previous studies (Montgomery & Sorkin, in press; Sorkin et al., 1991). The consistency of individual performance suggests that the individual members performed during the group trials at their measured performance (d') levels, thus allowing us to predict group performance (d'_{grp}) using the ideal model. However, there was no way to completely account for day-to-day variability and differences in individual performance during the group trials. A secondary finding was that individual differences were small, suggesting that there would be little benefit for the groups to differentiate individual performance during equal σ group interactions.

An important finding is that actual group performance (d'_{grp}) increased with each additional group member, as

predicted by the ideal equation. However, there were decreasing benefits with each additional group member; six and seven-member groups showed very small improvements over four member groups.

Another crucial factor in the group decision process is the correlation between group members' estimates (r). Correlation between informational sources leads to a reduction in the variability across sources and therefore, possible group performance. Group performance, under the correlated condition, was less than the independent condition as can be seen in Figure 8. The improvement in performance gained by each additional group member, during the correlated condition, was less than the independent condition. In a sense, by displaying correlated informational sources, we were not adding additional group members but partial performers.

An enlightening example of how high correlations between group members can lead to decreased group performance can be shown in an aircraft mishap. For example, when the members of an Air Florida flight crew all concentrated on the same single display gauge their input data was essentially perfectly correlated. As a consequence, key data from several other independent sources were ignored. They unknowingly crashed a landable airplane in the Florida Everglades, killing everyone on board, partly due to a increase in inter-member correlation.

As our results suggest and previously suggested by Hogarth (1978), it may be beneficial to add group members that lower the r value. Moreover, when making selections for decision making groups, a "yes" woman/man is not the best addition to the group. Someone who is not afraid to say what she/he really believes can lower the inter-member correlation and raise group performance.

The groups appeared to differentiate the differences between individual members during conditions where the variability of member performance was large. However, the individual weights appeared to be random when the variability of member performance was small. Prior research on group decision making suggests that groups tend to treat all information sources "as if" they were equal and disregard differences in information (Wickens, 1989). Our results in the unequal difficulty (σ) trials are an exception to these findings. The unequal σ condition showed that the groups could differentiate and effectively weight the sources of information. The group assigned more weight to information from the more reliable members and did not treat the information "as if" it were equal.

In the equal σ conditions, the groups may have reasoned that small individual differences in performance ability were not worth evaluating. The groups may have chosen to satisfice and speed up the interactions, thereby receiving most of the payoffs with less cognitive workload. Humans

have been shown to be heuristical; looking for cognitive shortcuts to answer questions quickly (Tversky & Kahneman, 1973).

The group efficiency measures showed that the overall group efficiency declined with the number of group members, and that the groups utilized a moderately effective weighting strategy. Figure 8, displaying the σ=3.0 condition, showed that the data could be fit well with a correction factor proportional to the number of group members (m). This correction factor may be in the form of an additional, additive noise that each group member adds to the group. Although it is improbable that each new group member adds an identical noise value, the correction factor suggests some important ideas about the best group size. Since the correction factor relates group size to actual group performance, it may be possible to define a desired level of performance and choose the corresponding size of the group. Likewise, the decreasing returns in group performance as the group became larger is an important finding.

The relatively high observed efficiency measures (η_{obs}) suggest that the model of the Ideal Group (Sorkin & Dai, 1994) is an effective predictor of group performance. The weighting efficiencies were consistently high across all groups, decreasing very little as m increased. Thus, the decreases in observed efficiency (η_{obs}) with additional group members cannot be explained by losses in weighting strategies.

The groups were able to assign weights ideally when the differences in individual performance were important. Nevertheless, even in the unequal σ condition there may have been some inappropriate weighting strategies. For example, weights may have been somewhat biased by social interaction characteristics during the unequal σ group trials. S_7 and S_8 were highly interactive, as was S_5. However, S_2 was reluctant to interact during the group trials and was reserved when she did speak during a trial as seen in Table 9. As Table 7 showed, the two unequal σ conditions where S_2 observed difficult ($\sigma = 3.0$) information during the trials produced larger d' values than when S_2 received less difficult ($\sigma = 1.5$) information. Even when they observed difficult information to discriminate, S_7 and S_8 received slightly more weight than ideal (as shown in Figure 13). In hindsight, it would have been beneficial to have a more objective measure of the individual subject tendencies than notes taken during the group interactions.

It is important to consider the additional weight assigned to the group member who was told to respond for the group during the trial. As Figure 17 showed, a member typically received more relative weight when they responded for the group than when they were just a part of the group. The responder would ideally be an unbiased integrator of the incoming information and respond according to the information transmitted by the group members. But this was not the case, and group responder received more weight.

This finding has significant implications on how group tasks are designed. The assignment of supervisory titles, as often is the case for the group responder, frequently carries more weight and could result in a non-ideal weighting strategy (Samuelson & Allison, 1994). Within a resource allocation task, Samuelson and Allison found that the assignment of supervisory titles to a member of the group led the "supervisor" to take more than their fair share of the resources. It is important for members of a group to realize that when they are to respond for the group, she/he tends to receive more weight than the other group members. If the group wants to perform ideally, they should avoid a bias to give their own judgment more weight.

In order to gain a better understanding of why observed efficiency decreased in larger groups, we need to examine some possible explanations. Individual performance (d'_i) may have decreased during group performance due to: social loafing, the free rider effect, the sucker effect (Shepperd, 1993), and a group resignation. When individuals become part of a group, there is a tendency for members of the group to loaf and neglect responsibility for the group's actions; Especially if the individual's efforts are unidentifiable (Latané, Williams, & Harkins, 1979); similar to the current study.

Shepperd's (1993) review paper offers an expectancy theory account of several reasons why individual effort and performance (d'_i) may have decreased during the group perf-

ormance. When an individual is put into a group and she/he believes her/his contribution to the overall goal is unimportant, she/he is tempted to free-ride on the coattails of the others in the group and still receive the group benefits (Olson, 1965; Kerr & Bruun, 1983; Williams & Karau, 1991). This may have occurred in the current study because the members of the group all receive the same monetary incentive for their performance and the individuals are free from separate evaluation (Harkins & Jackson, 1985).

Additionally, observations made by the experimenter during the difficult (σ=3.0) group trials showed an almost helpless attitude of some group members. When the group task became very difficult and there was no clear group decision, several of the group members tended to loaf and not interact fully. This perceived sense of resignation (learned helplessness) may have affected the group results and could be detrimental to group performance in other tasks.

Another loss in individual effort may have occurred because the threat of free-riding was present. Individuals may have reduced their effort in order to avoid being taken as a "sucker" by the free-riders in the group (Orbell & Dawes, 1981). An individual who feels she/he is being used to perform more than her/his fair share of the work, while others benefit, may reduce her/his own effort.

Shepperd (1993) examined an additional problem that the groups may have encountered in the current study. The

decreasing benefits of each new group member might be explained by a coordination loss. When the size of the group (m) increased, the ability of the group to perform effectively may have decreased due to a loss in coordination. An example of a loss due to coordination can be seen in a game of "tug-of-war". When there are two members in a group they can effectively communicate and pull in the same direction, at precisely the same moment. However, as the tug-of-war group grows in size it becomes more difficult for the group to effectively communicate. Some members may pull at different times and in slightly different directions.

There are several important practical and theoretical implications that stem from the current study. These findings may be helpful in increasing actual group performance and in understanding group decision processes in the future.

In order to gain a better understanding of the actual group performance, the variation of individual performance across conditions needs to be examined. It would not be a simple change, but it may be beneficial in the future to monitor individual performance simultaneously with group performance. This could also act as a deterrent to social loafing.

Within the context of the current task description, I think there is a point at which the benefits of additional group members are negated by the additional costs and problems that each new member adds to the group decision process. Tradeoffs associated with large groups involve:

additional noise, coordination loss and declining benefits with each member against the small performance increase each group member adds to the group. I suggest a compromise of four group members. Four members would be an excellent combination of group performance and the reduced efficiencies associated with the six and seven-member groups. Of course, the cost of making a mistake should play a part in how many group members are included in a decision making group. As our results demonstrate, if the cost of a mistake is great, then the more group members the better.

The groups may have assigned weights partially based on a non-ideal strategy. Several observations were made that suggested loudness or latency of response were used as weighting cues during some group interactions. There could be an effective way to conduct future research by measuring social interaction characteristics in order to gain a more complete understanding of the group weighting process.

Some suggestions for training effective decision making groups also follow from our current study. Groups need to be taught that social loafing and coordination loss are likely to occur, especially as group size increases, and to be alert for these problems. Another valuable lesson would be to train individuals to recognize personality differences but to concentrate on actual performance abilities. Loudness and affinity for interaction are not always the best performance indicators.

Similarly, groups need to be aware that the group responder typically receives more weight during the group interaction. If the group response can not be randomly distributed between the group members, then the group responder needs to be an efficient integrator of information and preferably the best performer in the group.

REFERENCES

Agresti, A. & Finlay, B. (1986). Statistical methods for the social sciences. San Francisco: Dellen Publishing Company.

Ashby, F. G. & Maddox, T. W. (1992). Complex decision rules in categorization: Contrasting novice and experienced performance. Journal of Experimental Psychology: Human Perception and Performance, 18, 50-71.

Berg, B. G. (1989). Analysis of weights in multiple observation tasks. Journal of the Acoustical Society of America, 86, 1743-1745.

Berg, B. G. (1990). Observer efficiency and weights in a multiple observation task. Journal of the Acoustical Society of America, 88, 149-158.

Berg, B. G. & Green, D. M. (1990). Spectral weights in profile listening. Journal of the Acoustical Society of America, 88, 758-766.

Bordley, R. F. (1986). One person/one vote is not efficient given information on factions. Theory and Decision, 21, 231-250.

Clement, D. E., Schiereck, J. J. Jr. (1973). Sex composition and group performance in a visual signal detection task. Memory & Cognition, 1, 251-255.

Condorcet, Nicolas Caritat de (1785). Essai sur l'application de l'analyse a la probabilite des decisions rendues a la pluralite des voix, Paris.

Durlach, N. I., Braida, L. D., & Ito, Y. (1986). Towards a model for discrimination of broadband signals. Journal of the Acoustical Society of America, 80, 60-72.

Fryman, J. (1994). Local access network interface: Novell interfacing from C+ to Turbo Pascal. [Computer Program]. Gainesville, FL: University of Florida, Psychology Computer Laboratory.

Gonzalez, R. (1994). When words speak louder than actions: Another's evaluations can appear more diagnostic than their decisions. Organizational Behavior and Human Decision Processes, 58, 214-245.

Green, D. M. (1992). The number of components in profile analysis tasks. Journal of the Acoustical Society of America, 91, 1616-1623.

Green, D. M. (1994). Signal detection theory revisited. Unpublished manuscript.

Green, D. M., & Swets, J. A. (1966). Signal detection theory. New York: Wiley.

Grofman, B., Feld, S. L., & Owen, G. (1984). Group size and the performance of a composite group majority: Statistical truths and empirical results. Organizational Behavior and Human Performance, 33, 350-359.

Grofman, B., Owen, G. & Feld, S. L. (1983). Thirteen theorems in search of the truth. Theory and Decision, 15, 261-278.

Harkins, S., & Jackson, J. (1985). The role of evaluation in eliminating social loafing. Personality and Social Psychology Bulletin, 11, 457-465.

Hinsz, V. B. (1990). Cognitive and consensus processes in group recognition memory performance. Journal of Personality and Social Psychology, 59, 705-718.

Hogarth, R. M. (1978). A note on aggregating opinions. Organizational Behavior and Human Performance, 21, 40-46.

Johnson, H. H. & Torcivia, J. H. (1967). Group and individual performance on a single-stage task as a function of distribution of individual performance. Journal of Experimental Psychology, 3, 266-273.

Kerr, N. L., & Bruun, S. E. (1983). Dispensability of member effort and group motivation losses: Free-rider effects. Journal of Personality and Social Psychology, 44, 78-94.

Latané, B., Williams, K. & Harkins, S. (1979). Many hands make light the work: The causes and consequences of social loafing. Journal of Personality and Social Psychology, 37, 822-832.

Lutfi, R. A. (in press). A correlational method for deriving weights in multiple-observation tasks. Journal of the Acoustical Society of America.

Montgomery, D. A. (1991). Information selection under processing limitations: A weight analysis. Master's thesis, University of Florida, Gainesville.

Montgomery, D. A. (1993). The effects of display coding factors on observer visual signal detection. Doctoral dissertation, University of Florida, Gainesville.

Olson, M. (1965). The logic of collective action: Public goods and the theory of groups. Cambridge, MA: Harvard University Press.

Orbell, J. & Dawes, R. (1981). Social dilemmas. In G. Stephenson & H.H. Davis (Eds.), Progress in applied psychology (Vol. 1, pp. 37-65). New York: Wiley.

Samuelson, C. D. & Allison, S. T. (1994). Cognitive factors affecting the use of social decision heuristics in resource sharing tasks. Organizational Behavior and Human Decision Processes, 58, 1-27.

Sanders, M. S. & McCormick, E. J. (1987). Human factors in engineering and design. (6th ed.) New York: McGraw-Hill.

Schum, D. A. (1975). The weighting of testimony in judicial proceedings from sources having reduced credibility. Human Factors, 17, 172-182.

Shepperd, J. A. (1993). Productivity loss in performance groups: A motivation analysis. Psychological Bulletin, 113, 67-81.

Sniezek, J. A. (1990). A comparison of techniques for judgmental forecasting by groups with common information. Group & Organizational Studies, 15, 5-19.

Sniezek, J. A., & Henry, R. A. (1989). Accuracy and confidence in group judgment. Organizational Behavior and Human Decision Processes, 43, 2-28.

Sorkin, R. D. & Dai, H. (1994). Signal detection analysis of the ideal group. Journal of Organizational Behavior and Human Decision Processes, 60, 1-13.

Sorkin, R. D., Mabry, T. R., Weldon, M., & Elvers, G. (1991). Integration of information from multiple element display. Organizational Behavior and Human Decision Processes, 49, 167-187.

Sorkin, R. D. & Montgomery, D. A. (1991). Effect of time compression and expansion on the discrimination of tonal patterns. Journal of Acoustical Society of America, 90, 846-857.

Sorkin, R. D., Robinson, D. E. & Berg, B. G. (1987). A detection theory method for the analysis of visual and auditory displays. Proceedings of the Human Factors Society 31st annual meeting, 1184-1187.

Tanner , W. P. & Birdsall, T. G. (1958). Definitions of d' and ^{n} as psychophysical measures. Journal of Acoustical Society of America, 30, 922-928.

Thorndike, R. M. (1978). Correlational procedures for research. New York: Gardner Press, Inc.

Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. Science, 185, 1124-1130.

Vollrath, D. A., Sheppard, B. H., Hinsz, V. B., Davis, J. H. (1989). Memory performance by decision-making groups and individuals. Organizational Behavior and Human Decision Processes, 43, 289-300.

Welkowitz, J., Ewen, R. B., & Cohen, J. (1982). Introductory statistics for the behavioral sciences. (3rd ed.) New York: Academic Press, Inc.

Wickens, C. D. (1984). Engineering psychology and human performance. Columbus, OH: Charles E. Merrill Publishing Co.

Williams, K. D., & Karau, S. J. (1991). Social loafing and social compensation: The effects of expectations of co-worker performance. Journal of Personality and Social Psychology, 61, 570-581.

APPENDIX
LUTFI'S CORRELATIONAL WEIGHTING ANALYSIS

It is enlightening to evaluate the relative weight (influence) of each group member during the group interaction and decision. Similarly, by evaluating deviations from ideal weights we may gain a better understanding of the inefficiencies in the group interaction. According to Sorkin and Dai's (1994) group model employing Signal Detection Theory (Green & Swets, 1966; Green, 1994), during the group decision process, the group assigns weights (a_i) to each individual's estimate of their observation (x_i) to reach a group decision statistic (Z).

The relative weight assigned to each individuals' observations in the current study was estimated using a correlational weighting technique (Lutfi, in press). Lutfi's analysis is similar to COSS (Conditional on a single stimulus) analysis suggested by Berg (1989), except that Lutfi's correlational analysis offers a least-squares estimate of weights instead of a maximum likelihood estimate. The correlational weighting analysis was used because it makes no assumptions about independence between observers.

Since the correlational method for deriving weights is based on some of the same principles as COSS analysis we will briefly review COSS. COSS analysis is a powerful tool that can be used to determine the weights assigned to

multiple, independent sources of information. This analysis enables the experimenter to evaluate the separate effects on an observer's response due to small changes in the values of independent elements over a block of trials. These effects on the observer's response, which are caused by the perturbations in the elements, are referred to as COSS functions. A COSS function is a psychometric function with the probability of responding signal plotted on the y-axis and the value of the informational source along the x-axis. In order to analyze the weights assigned to each source of information, a COSS function is determined for each element independent of all of the other elements in the trial. By measuring the slope of these COSS functions, a maximum likelihood estimate for each of the relative weights given to each element of a multi-element display can be determined. Therefore, the steeper the slope of a COSS function, the more weight given to that particular element in the display. Furthermore, Berg (1989) proposed that the weights evaluated using COSS analysis are not biased by the information content of the elements or by additive sources of internal noise.

The COSS analysis has demonstrated to be beneficial in several research areas. For instance, it was used in a hearing detection task (Berg & Green, 1990) to determine which channel is given the most weight. It was also utilized in a short duration visual detection task (Sorkin et al., 1991) to determine which gauges are given the most

weight. Furthermore, COSS estimates were used to show that observers can use emergent cues to correctly weight information in a multiple gauge, visual detection task (Montgomery & Sorkin, in press). However, the assumption of independent observations that COSS analysis requires will fail in the current study (r=0.25). It was for this reason that the correlational method of weighting will be employed.

Correlational Weighting Analysis

Assume that three individuals view separate displays made up of nine gauges as shown in Figure 3, and interact freely to make a group decision. Group members 1, 2, and 3 are presented stimulus values that are normally distributed and are statistically independent with variance σ^2_{x1}, σ^2_{x2}, and σ^2_{x3}.

The decision process for the group is demonstrated in Figure 1. The decision rule assumes that the group's response is determined by the value of a decision variable, Z. The group decision variable, Z, is based on a weighted sum of the estimates of observations made by group members 1, 2, and 3 (x_1 , x_2, and x_3) with error (ε). This value is shown for the group in equation A1.

$$Z= a_1x_1 + a_2x_2 + a_3x_3 + \varepsilon \tag{A1}$$

The variables a_1, a_2, and a_3 represent the relative weights assigned to the estimates made by group members 1, 2, and 3 respectively, and ε is the sum of all additive errors. This error may be due to noise in the information displays or

additional noise during the interaction process. Since x_1, x_2, and x_3 are assumed to be independent, equation A1 is the general equation for multiple linear regression. (Agresti & Finlay, 1986, p.249-251). Therefore, Z is the dependent variable and x_1, x_2, and x_3 are the linear predictor variables in equation A1. The total variance in the dependent variable, Z, can be separated into four parts: the four parts would include Z's relation to x_1, x_2, and x_3, respectively, and the variance that is left over. (Lutfi, in press)

$$\sigma^2_Z = a_1{}^2\sigma^2_{x1} + a_2{}^2\sigma^2_{x2} + a_3{}^2\sigma^2_{x3} + \sigma^2_\varepsilon \qquad \text{(A2)}$$

Equation A2 can be rewritten by expressing each variable as a proportion of the total variance accounted for,

$$1 = \rho^2_{Zx1} + \rho^2_{Zx2} + \rho^2_{Zx3} + \rho^2_{Z\varepsilon} \,. \qquad \text{(A3)}$$

The correlation coefficients ρ_{Zx1}, ρ_{Zx2}, and ρ_{Zx3} can then be calculated by using linear least-squares regression of Z on x_1, Z on x_2, and Z on x_3.

The weights that we are estimating correspond closely to the regression coefficients in the analysis above. These relative weights are computed by the formulas $a_1 = \rho_{Zx1}\,\sigma_Z / \sigma_{x1}$, $a_2 = \rho_{Zx2}\,\sigma_Z / \sigma_{x2}$, and $a_3 = \rho_{Zx3}\,\sigma_Z / \sigma_{x3}$. Conse-

quently, these equations can be used to form an equation for the relative weights:

$$\frac{a_1}{a_2} = \frac{\rho_{Zx1}\ \sigma_{x2}}{\rho_{Zx2}\ \sigma_{x1}}$$

$$\frac{a_1}{a_3} = \frac{\rho_{Zx1}\ \sigma_{x3}}{\rho_{Zx3}\ \sigma_{x1}} \qquad \text{(A4)}$$

Now, all that is required to solve for the relative weights are the estimates for ρ_{Zx1}, ρ_{Zx2} and ρ_{Zx3}.

Since Z in the current study corresponds to the group's subjective decision variable, it is impossible to obtain direct estimates of ρ_{Zx1}, ρ_{Zx2} and ρ_{Zx2}. Therefore, we are forced to estimate ρ_{Zx1}, $\rho_{Zx2}$and ρ_{Zx3} from the group response's relation to each group member's observation (sum of stimulus values). Response (R) is either signal (R_1) or noise (R_0), and therefore is a dichotomous variable. During the group trials, we assume the individual observers base their estimates (x_i) on the values of the stimuli on the vertical gauges. Based on this assumption the sum of the stimulus values for each group member (s_i) was used to estimate the influence given to each individual member. We calculated the point-biserial correlations, r_{Rx1}, r_{Rx2} and r_{Rx3}, between the group response and the observers' sum of elements (s_i) to estimate ρ_{Zx1}, ρ_{Zx2} and ρ_{Zx3}. These correlations are standard product-moment correlations and can be computed from standard formulas, over a block of trials,

i.e.

$$r_{Rxi} = \frac{n\Sigma Rs_i - \Sigma R\Sigma s_i}{\sqrt{n\Sigma R^2 - (\Sigma R)^2}\sqrt{n\Sigma s_i{}^2 - (\Sigma s_i)^2}} \qquad \text{(A5)}$$

The correlational formula for the relative weights are then:

$$\frac{a_1}{a_2} = \frac{r_{Rx1}\ \sigma_{x2}}{r_{Rx2}\ \sigma_{x1}}$$

$$\frac{a_1}{a_3} = \frac{r_{Rx1}\ \sigma_{x3}}{r_{Rx3}\ \sigma_{x1}} \qquad \text{(A6)}$$

Where r_{Rx1}, r_{Rx2}, and r_{Rx3} are the point biserial correlations between the sum of the stimulus values presented to each group member (s_i) and the group response (R). σ_{x1}, σ_{x2}, and σ_{x3} are the standard deviations of the stimulus distributions presented to each observer and represent the difficulty of information. To estimate the relative weights for larger groups equation A6 above is extended in sequence. To solve for the actual relative weights, we assume relative weight $a_1 = 1$ and $\sigma_i = \sigma_i$ for all i (equal difficulty distribution case). The two equations have three weights, therefore by arbitrarily assigning $a_1 = 1$, the relative weights can be solved. The relative weights are then normalized to allow comparisons between conditions. In order to normalize the weights, each individual weight is divided by the sum of all the individual weights.

An example is given below to show how Lutfi's correlational analysis is used to derive the relative weights. A group of four individuals receive independent, equal σ

information. Again, we assume that the subjects followed instructions and are using the stimuli values to make their estimates. Therefore, the sum of the stimulus values would provide a good estimate of each individual's observation. The individual then conveys information about her/his observation to the group during the group interaction. r_{Rx1} represents the point-biserial correlation between group member 1's sum of the stimulus displays (s_1) and the group's response (R), over a block of 100 trials. In order to determine the normalized relative weights, these correlations along with the σ_i (difficulty of each of the subject's information) are substituted into equation A6. The standard deviation of a single stimulus display gauge (σ_i), is referred to as the difficulty of the information. In order to determine the relative weights, the standard deviation of the entire stimulus display screen needs to be calculated; for a sum of normal values, this is $\sigma_i/\sqrt{n}$, where n is the number of stimulus gauges.

An example of the group data, with four members, shows:

$r_{Rx1} = .37$	$\sigma_i = 1.5$ for all σ_i
$r_{Rx2} = .23$	standard deviation of the
$r_{Rx3} = .19$	stimulus display = 1.5 / 3
$r_{Rx4} = .24.$	=0.50

Substituting these values into the relative weights equation above would give the following values,

$$\frac{a_1}{a_2} = \frac{r_{Rx1}\sigma_2}{r_{Rx2}\sigma_1} = \frac{(.37)*(.5)}{(.23)*(.5)} = 1.609$$

$$\frac{a_1}{a_3} = \frac{r_{Rx1}\sigma_3}{r_{Rx3}\sigma_1} = \frac{(.37)*(.5)}{(.19)*(.5)} = 1.947$$

$$\frac{a_1}{a_4} = \frac{r_{Rx1}\sigma_4}{r_{Rx4}\sigma_1} = \frac{(.37)*(.5)}{(.24)*(.5)}. = 1.542$$

After substituting $a_1 = 1$ into the equations above the relative weights will be:

$a_1 = 1$, $a_2 = .622$, $a_3 = .514$, $a_4 = .649$

In order to normalize the weights, we divide each relative weight above by the sum of the weights, then

$a_1 = .36$, $a_2 = .22$, $a_3 = .18$, and $a_4 = .23$.

From this example, we see that group member 1 has a higher relative weight than the other observers during the interaction. Consequently, we conclude that during the group interaction, the group assigns the most influence to what member 1 says and the least influence to what member 3 says.

<u>Statistical Significance of Weights</u>

The correlational method of estimating the relative weights given to different sources of information is based on a multiple regression analysis of trial-by-trial data. This method of analysis makes it possible to do statistical tests on the relative weights assigned to each observer. Tests of statistical significance can be performed using standardized Z-scores. The expression for the point-biserial correlation (equation A5) is very similar to the expres-

sion for the standardized difference between means.

$$r_{Rxi} = \frac{\bar{s}_{i(0)} - \bar{s}_{i(1)}}{\sigma_{si}} \sqrt{p_0 p_1} \qquad \text{(A7)}$$

where $\bar{s}_{i(0)}$ ($\bar{s}_{i(1)}$) and p_0 (p_1) are, respectively, the mean value of s_i and the proportion of trials for which the response was noise (R_0) and signal (R_1). The test for the significance of r_{Rxi}, and therefore the weight, is

$$Z = \frac{r_{Rxi}}{1 / \sqrt{n}} \qquad \text{(A8)}$$

where n>30 trials, $1 / \sqrt{n}$ may be used as the standard error of r_{Rxi}. Using equation A8, we can test the significance of correlations between each observer's sum of elements and the group response (R). It can be determined if the weights are statistically different from one another by substituting the difference in correlations between observers (r_{Rx1} - r_{Rx2}) into the numerator of equation A8. Equation A8 can also be used to determine confidence intervals for the weights.

Another benefit of the correlational weighting analysis is that it can account for non-independent observations. If the inter-member correlation (r)>> 0, then the partial correlation ($r_{Rx1 \cdot x2}$) can be substituted into equation A6 to determine the relative weights. The partial correlation measures the correlation between an observation and the group response while accounting for the other observations' relationship to the group response. In groups with more

than two members, higher order partial correlations need to be used ($r_{Rx_1 \cdot x2,\ x3,\ x4}$).

During the current study, the inter-member correlations (r) were either experimentally set to 0.0 or 0.25. The actual inter-member correlations (between observations) were analyzed and checked for statistical significance at the p value (two-tail) = .01 on all conditions. All actual inter-member correlations in the r=0.25 experimental condition were not statistically different from 0.25. The relative weights for these conditions were then evaluated using partial correlations. Additionally, due to random error it is likely that r >> 0.0 for some of the r=0.0 experimental conditions. The actual correlations in this condition were tested at the same p value. Out of all the possible correlations, only four were found to significantly different from zero. Based on the large number of correlations and a Monte Carlo simulation, these were assumed to be due to random error and partial correlations were not utilized in the r=0.0 condition.

The partial correlations can be tested for significance by adjusting the standard error used in equation A8. In the first order partials (group of 2 members) the denominator of equation A8 becomes $1/\sqrt{N-4}$. For second and higher order partial correlations the appropriate test for significance of the correlation becomes:

$$Z = \frac{r_{Rx1 \cdot x2, x3, \ldots xm}}{1/\sqrt{N - (m + 3)}} \tag{A9}$$

where m is the number of variables partialled out (Thorndike, 1978).

BIOGRAPHICAL SKETCH

Christopher J. Hays was born March 31, 1970. He received his Bachelor of Science degree in behavioral sciences and officer commission from the United States Air Force Academy in May, 1992. He received his pilot wings in July, 1993. He received his M.S. from the University of Florida (1995). Christopher's research interests are focused on human perception, group interaction, and human factors within the cockpit.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Robert D. Sorkin

Robert D. Sorkin, Chair
Professor of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

David M. Green

David M. Green
Graduate Research Professor of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

James Shepperd

James Shepperd
Assistant Professor of Psychology

This thesis was submitted to the Graduate Faculty of the Department of Psychology in the College of Liberal Arts and Sciences and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Master of Science.

May, 1995

Dean, Graduate School